NBER WORKING PAPER SERIES

THE TERM STRUCTURE OF CURRENCY CARRY TRADE RISK PREMIA

Hanno Lustig
Andreas Stathopoulos
Adrien Verdelhan

Working Paper 19623
http://www.nber.org/papers/w19623

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2013, Revised October 2017

Many thanks to Robert Barro, Mikhail Chernov, Riccardo Colacito, Ian Dew-Becker, Greg Duffee, Bernard Dumas, Emmanuel Farhi, Jon Faust, Ron Giammarino, Stefano Giglio, Lars Hansen, Espen Henriksen, Urban Jermann, Leonid Kogan, Karen Lewis, Matteo Maggiori, Ian Martin, Stefan Nagel, Jonathan Parker, Tarun Ramadorai, Robert Ready, Lucio Sarno, Jose Scheinkman, Alan Taylor, Andrea Vedolin, Mungo Wilson, Jonathan Wright, Fernando Zapatero, Irina Zviadadze, seminar participants at the Federal Reserve Board, Georgetown University, Harvard, LSE, LBS, MIT, UC3 in Madrid, UC Davis, the Said School at Oxford, Cass at City University London, Syracuse University, University of Bristol, UBC, University of Exeter, University of Lausanne, University of Massachusetts, University of Michigan, University of Rochester, USC, and Wharton, as well as the participants at the First Annual Conference on Foreign Exchange Markets at the Imperial College, London, the International Macro Finance Conference at Chicago Booth, the Duke/UNC Asset Pricing Conference, the WFA Meetings, and the NBER Summer Institute. A previous version of this paper circulated under the title “Nominal Exchange Rate Stationarity and Long-Term Bond Returns.” The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Term Structure of Currency Carry Trade Risk Premia
Hanno Lustig, Andreas Stathopoulos, and Adrien Verdelhan
NBER Working Paper No. 19623
November 2013, Revised October 2017
JEL No. F20,F31,G12

ABSTRACT

Fixing the investment horizon, the returns to currency carry trades decrease as the maturity of the foreign bonds increases, because the local currency term premia offset the currency risk premia. The time series predictability of foreign bond returns in dollars similarly declines as the maturity of the bonds increases. Leading no-arbitrage models in international finance cannot match the downward term structure of currency carry trade risk premia. While currency risk premia on short-term bonds reflect differences in transitory and permanent risk, we show that the premia on long-term bonds only reflect differences in the risk of permanent shocks to investors' marginal utility.

Hanno Lustig
Stanford Graduate School of Business
655 Knight Way
Stanford, CA 94305
and NBER
hlustig@stanford.edu

Adrien Verdelhan
MIT Sloan School of Management
100 Main Street, E62-621
Cambridge, MA 02142
and NBER
adrienv@mit.edu

Andreas Stathopoulos
Finance Department
University of Wahington
astathop@marshall.usc.edu

The uncovered interest rate parity (U.I.P.) condition states that expected exchange rate changes should be exactly offset by interest rate differences across countries, so expected currency excess returns should be always equal to zero. The U.I.P. condition fails empirically in the short run: a U.S. investor who trades foreign Treasury bills and does not hedge her currency risk exposure earns predictable dollar excess returns that depend on ex ante observable currency characteristics, such as the country's interest rate (Hansen and Hodrick, 1980; Fama, 1984) and the slope of the yield curve (Ang and Chen, 2010). The profitability of the classic currency carry trade, which goes long in Treasury bills of high interest rate currencies and short in low interest rate currencies, is a well-known example (Lustig and Verdelhan, 2007).

Our paper explores the properties of the same investment strategy implemented with long-term bonds, and compares it to the standard strategy that uses Treasury bills. In particular, we fix the holding period at one month, we focus on the same set of G10 countries, and we consider the most important predictors of bond and currency returns: the level and slope of the yield curve. The first strategy we consider goes long the bonds of high interest rate strategies and short the bonds of low interest rate strategies, whereas the second strategy goes long the bonds of flat yield curve currencies and short the bonds of steep yield curve currencies. We jointly refer to those two strategies as the carry trade strategy. It is key to note that the investment horizon is one-month, as in the classic tests of the U.I.P. condition, not ten years or more as in the tests of the U.I.P condition over long horizons. We find that, as the maturity of the bonds increases, the average excess return declines to zero. In other words, whereas the carry trade implemented with Treasury bills is profitable, the carry trade implemented with long-term bonds is not. Similar results emerge in individual country time-series predictability tests: as the maturity of the bonds increases, the predictability of the cross-country differences in dollar bond returns disappears.

The downward-sloping term structure of carry trade risk premia that we uncover is a challenge for the leading models in international finance. While some recent no-arbitrage models replicate the U.I.P. puzzle, they fail to replicate the carry trade risk premia at the long end of the yield curve. To illustrate this point, we simulate the multi-country model of Lustig, Roussanov, and Verdelhan (2011) and show that it implies a counterfactual flat term structure of currency carry trade risk premia. Therefore, our empirical findings on the long-term bond returns expressed in the same currency raise the bar for models in international finance and our paper explains why by deriving a preference-free condition that dynamic asset pricing models need to satisfy.

In particular, we show that the difference between domestic and foreign long-term bond risk premia, expressed in domestic currency, is determined by the difference in the volatilities of the permanent components of the stochastic discount factors (SDFs). Therefore, to obtain zero currency carry trade risk premia at the long end of the yield curve, the conditional risk in the permanent components needs to be equalized across currencies – in other words, the shocks to the quantity or price of risk cannot impact marginal utility differences, and hence exchange rates, in the long run. We refer to this condition as the long-run risk neutrality of exchange rates. Armed with our preference-free condition, we revisit a large class of dynamic asset pricing models that have been used to study U.I.P. violations, ranging from the Campbell and Cochrane (1999) external habit model, the Bansal and Yaron (2004) long run risks (LRR) model, the disaster risk model, and the reduced-form model of Lustig, Roussanov, and Verdelhan (2011). In each case, we derive the restrictions on the model's parameters that are necessary in order to match our empirical findings, sometimes leading to degenerate model solutions.

Our preference-free condition should help design the next generation of models because the downward-sloping term structure of carry trade risk premia is informative about the kind of risks that drive exchange rates and bond returns and the properties of SDFs. The finance literature has established that SDFs need to be highly volatile (see Hansen and Jagannathan, 1991) and that most of the risk needs to be permanent (see Alvarez and Jermann, 2005; Hansen and Scheinkman, 2009). Moreover, to replicate the failure of U.I.P. in no-arbitrage models, the SDFs need to be more volatile in countries with currently low interest rates (see Backus, Foresi, and Telmer, 2001). Finally, to generate profitable currency carry trades, the SDFs need to be more exposed to global risk in low interest rate countries (see Lustig, Roussanov, and Verdelhan, 2011). We contribute to this literature by showing that our empirical findings impute a key role to transitory risk in currency markets, since the permanent SDF components should exhibit the same volatilities. We also derive a lower bound on the correlation between the domestic and foreign permanent SDF components. The lower bound depends on the maximum risk premium in the domestic and foreign economies, the domestic and foreign term premia, as well as the volatility of the permanent component of the exchange rate changes. The large equity risk premium, the low term premium, and the low volatility of the permanent component of the exchange rate changes imply a lower bound of at least 0.90. While Brandt, Cochrane, and Santa-Clara (2006) show that SDFs are highly correlated across countries, we show that this high correlation comes from the permanent SDF components, which are the key drivers of equity risk premia and the welfare cost of the business cycles (see

Alvarez and Jermann, 2004).

Finally, our theoretical framework sheds light on the long-run U.I.P condition, which compares foreign and domestic long-term interest rates to long-term changes in exchange rates. Meredith and Chinn (2005) find that long-run U.I.P is a potentially valid description of the data. However, empirical tests lack power in finite samples: intuitively, there are few non-overlapping observations of 10-year windows available so far. We show that long-run U.I.P. holds on average when exchange rates are not subject to permanent shocks, and thus are stationary in levels. In the absence of permanent shocks, countries with high long-term interest rates have stronger currencies (the level of the exchange rate is temporarily high), so their currencies are expected to depreciate in the long run (see Dornbusch, 1976; Frankel, 1979, for early contributions on the relation between the level of the exchange rate and interest rates). A similar reasoning applies to Engel's (2016) puzzling finding that currencies with high short-term interest rates earn high excess returns (and hence are risky) in the short run, but these same currencies produce low excess returns (and hence are not risky) in the long run.

Our analysis is subject to two important caveats: we proceed under the assumption that financial markets are complete and that long-term bond returns can be approximated in practice by 10 and 15-year bond returns. The first assumption is certainly counterfactual but provides a natural benchmark. We show that imposing the absence of arbitrage only on Treasury bills generalizes our results to a large class of incomplete market models that feature constant relative risk-aversion and heteroskedastic consumption growth. The second assumption is supported by the simulation of the state-of-the-art Joslin, Singleton, and Zhu (2011) term structure model. Moreover, Martin and Ross (2013) show that abstracting from the martingale component does not come at a great cost when pricing domestic bonds and that the returns on long-term bonds converge to those of infinite maturity bond returns at an exponent rate in a Markov environment without a martingale component in the pricing kernel. Our preference-free condition does not apply to exchange rate models that abandon complete markets (see, e.g., Gabaix and Maggiori, 2015, and Bacchetta and van Wincoop, 2005, for leading examples). It remains to be determined whether these models can fit our facts, so we leave this as an open question for future research.

The rest of the paper is organized as follows. Section 1 rapidly reviews the literature. Section 2 focuses on the time-series and cross-section of foreign bond risk premia. Section 3 compares recent no-arbitrage models to the empirical term structure of currency carry trade risk premia. In Section 4, we derive the no-arbitrage, preference-free theoretical restriction imposed on bond returns and SDFs. Section 5 links long-term U.I.P. to exchange rate stationarity. Section 6 uses the previous theoretical results to analyze a large set of international finance models. Section 7 derives a lower bound on the correlation of the permanents components of SDFs and proposes a new benchmark to characterize bond returns across countries. In Section 8, we present concluding remarks. An Online Appendix contains supplementary material and all proofs not presented in the main body of the paper.

1 Related Literature

Our paper is related to three large strands of the literature: the literature on U.I.P. in the short run and the long run, the literature on empirical term premia across countries, and the literature that studies the general properties of SDFs. Our paper builds on the vast literature on the short-run UIP condition and the currency carry trade (see e.g. Engel, 1996, and Lewis, 2011, for recent surveys). Lustig, Roussanov, and Verdelhan (2011) focus on accounting for short-run uncovered interest rate parity condition (UIP) deviations and short-term carry trades respectively. They show that asymmetric exposure to global innovations to the pricing kernel are key to explaining the global currency carry trade premium at short maturities in a large class of dynamic term structure models.

Our work shows that these global innovations cannot have long-run effects on investors' marginal utility. In other words, if the carry trade risk premium is zero at the long end of the yield curve, then countries cannot differ in their exposure to permanent risk. Thus, in a large class of no-arbitrage models, carry trade risk premia at the short end of the yield curve simply compensates investors for bearing temporary shocks to their marginal utilities of wealth. Turning to long-run U.I.P., we are the first to derive general conditions under which long-run unconditional UIP holds.

Our focus is on the cross-sectional and time-series relation between the slope of the yield curve, interest rates, and exchange rates. We study whether investors earn higher returns on foreign bonds from countries in which the slope of the yield curve is higher than the cross-country average. Prior work, from Campbell and Shiller (1991) to Bekaert and Hodrick (2001) and Bekaert, Wei, and Xing (2007), focus mostly on the time series, testing whether investors earn higher returns on foreign bonds from a country in which the slope of the yield curve is currently higher than average for that country. Our results are consistent with, but not identical to the Campbell and Shiller (1991)-type time-series findings. There is no mechanical link between the time-series evidence and our cross-sectional result on the relative magnitudes of currency and bond risk premia. Time-series regressions test whether a predictor that is higher than its average implies higher returns, while the cross-sectional tests show whether a predictor that is higher in one country than in others implies higher returns in that country. Chinn and Meredith (2004) document some time-series evidence that supports a conditional version of UIP at longer holding periods, while Boudoukh, Richardson, and Whitelaw (2016) show that past forward rate differences predict future changes in exchange rates. Some papers study the cross-section of bond returns: Koijen, Moskowitz, Pedersen, and Vrugt (2016) and Wu (2012) examine the currency-hedged returns on 'carry' portfolios of international bonds, sorted by a proxy for the carry on long-term bonds, but they do not examine the interaction between currency and term risk premia, the topic of our paper. Ang and Chen (2010) and Berge, Jordà, and Taylor (2011) have shown that yield curve variables can also be used to forecast currency excess returns. These authors do not examine the returns on foreign bond portfolios.[1]

We interpret our empirical findings using a preference-free decomposition of the pricing kernel, building on recent work in the exchange rate and term structure literatures. On the one hand, at the short end of the maturity curve, currency risk premia are high when there is less overall risk in foreign countries' pricing kernels than at home (Bekaert, 1996, Bansal, 1997, and Backus, Foresi, and Telmer, 2001). On the other hand, at the long end of the maturity curve, local bond term premia compensate investors mostly for the risk associated with transitory innovations to the pricing kernel (Bansal and Lehmann, 1997; Hansen and Scheinkman, 2009; Alvarez and Jermann, 2005; Hansen, 2012; Hansen, Heaton, and Li, 2008; Bakshi and Chabi-Yo, 2012; Backus, Boyarchenko, and Chernov, 2016). Backus, Gregory, and Zin (1989) offer for one of the earliest analyses of long run bond yields and forward rates. In this paper, we combine those two insights to derive preference-free theoretical results under the assumption of complete financial markets. Foreign bond returns allow us to compare the permanent components of the SDFs, which as Alvarez and Jermann (2005) show, are by far the main drivers of the SDFs.

[1] Dahlquist and Hasseltoft (2013) study international bond risk premia in an affine asset pricing model and find evidence for local and global risk factors. Jotikasthira, Le, and Lundblad (2015) study the co-movement of foreign bond yields through the lenses of an affine term structure model. Our paper revisits the empirical evidence on bond returns without committing to a specific term structure model.

2 Foreign Bond Returns in the Time-Series and Cross-Section

We first describe the data and the notation, and then turn to our empirical results on the time-series and cross-sectional properties of foreign government bond returns.

2.1 Data

Our benchmark sample, to which we refer as the G10 sample, consists of a small homogeneous panel of developed countries with liquid bond markets: Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, and the U.K. The domestic country is the United States. For those countries, we calculate the returns of both coupon and zero-coupon bonds.

Our data on total return bond indices come from Global Financial Data. The dataset includes a 10-year government bond total return index, as well as a Treasury bill total return index, in U.S. dollars and in local currency, for each of our target countries. The data are monthly, starting in 1/1951 and ending in 12/2015. We use the 10-year bond returns as a proxy for long-maturity (theoretically infinity-approaching maturity) bond returns. While Global Financial Data offers, to the best of our knowledge, the longest time-series of government bond returns available, the series have two key limitations. First, they pertain to coupon bonds, while the theory presented in this paper pertains to zero-coupon bonds. Second, they only offer 10-year bond returns, not the entire term structure of bond returns. To address these issues, we also use zero-coupon bond prices.

Our zero-coupon bond dataset covers the same benchmark sample of G10 countries, but from at most 1/1975 to 12/2015. We use the entirety of the dataset in Wright (2011) and complement the sample, as needed, with sovereign zero-coupon curve data sourced from Bloomberg, estimated from government notes and bonds as well as interest rate swaps of different maturities.[2] Yields are available for bond maturities ranging from three months to 15 years, in three-month increments.

Finally, we collect data on inflation rates and sovereign credit ratings. Inflation rates are calculated using monthly Consumer Price Index (CPI) data from Global Financial Data, whereas sovereign credit ratings are

[2] The panel is unbalanced: for each currency, the sample starts with the beginning of the Wright (2011) dataset. The starting dates for each country are as follows: 2/1987 for Australia, 1/1986 for Canada, 1/1973 for Germany, 1/1985 for Japan, 1/1990 for New Zealand, 1/1998 for Norway, 12/1992 for Sweden, 1/1988 for Switzerland, 1/1979 for the U.K., and 12/1971 for the U.S. For New Zealand, the data for maturities above 10 years start in 12/1994.

from Standard & Poor's, available over the 7/1989–12/2015 period.[3] In order to construct credit rating averages for portfolios formed before 7/1989, we backfill each country's credit rating by assuming that the country's rating before 7/1989 is equal to its rating at the first available observation.

2.2 Notation

We now introduce our notation for bond prices, exchange rates, and bond and currency returns. In all cases, foreign variables are denoted as the starred version of their U.S. counterpart.

Bonds $P_t^{(k)}$ denotes the price at date t of a zero-coupon bond of maturity k, while $y_t^{(k)}$ denotes its continuously compounded yield: $\log P_t^{(k)} = -ky_t^{(k)}$. The one-period holding return on the zero-coupon bond is $R_{t+1}^{(k)} = P_{t+1}^{(k-1)}/P_t^{(k)}$. The log excess return on the domestic zero-coupon bond, denoted $rx_{t+1}^{(k)}$, is equal to

$$rx_{t+1}^{(k)} = \log\left[R_{t+1}^{(k)}/R_t^f\right], \tag{1}$$

where the risk-free rate is $R_t^f = R_{t+1}^{(1)} = 1/P_t^{(1)}$. Finally, r_t^f denotes the log risk-free rate: $r_t^f = \log R_t^f = y_t^{(1)}$.

Exchange Rates The nominal spot exchange rate in foreign currency per U.S. dollar is denoted S_t. Thus, an increase in S_t implies an appreciation of the U.S. dollar relative to the foreign currency. The log currency excess return is given by

$$rx_{t+1}^{FX} = \log\left[\frac{S_t}{S_{t+1}}\frac{R_t^{f,*}}{R_t^f}\right] = r_t^{f,*} - r_t^f - \Delta s_{t+1}, \tag{2}$$

and is the log excess return of a strategy in which the investor borrows at the domestic risk-free rate, R_t^f, invests at the foreign risk-free rate, $R_t^{f,*}$, and converts the proceeds into U.S. dollars at the end of the period.

Bond Risk Premia The log return on a foreign bond position (expressed in U.S. dollars) in excess of the domestic (i.e., U.S.) risk-free rate is denoted $rx_{t+1}^{(k),\$}$. It can be expressed as the sum of the bond log excess

[3] To construct averages of credit ratings, we assign each rating to a number, with a smaller number corresponding to a higher rating. In particular, a credit rating of AAA corresponds to a numerical value of 1, with each immediately lower rating getting assigned the immediately higher numerical value: AA+ corresponds to a numerical value of 2 and AA to 3, all the way down to CC- (22) and SD (23). We also construct rating series adjusted for outlook, as follows: a 'Negative' outlook corresponds to an upward adjustment of 0.5 in the numerical value of the rating, a 'Watch Negative' outlook to an upward adjustment of 0.25, a 'Stable' or 'Satisfactory' outlook to no adjustment, and a 'Positive', 'Strong' or 'Very Strong' outlook to a downward adjustment of 0.5. For example, a credit rating of BB (coded as 12) receives a numerical value of 12.5 with a 'Negative' outlook and a value of 11.5 with a 'Positive' outlook.

return in local currency plus the log excess return on a long position in foreign currency:

$$rx_{t+1}^{(k),\$} = \log\left[\frac{R_{t+1}^{(k),*}}{R_t^f}\frac{S_t}{S_{t+1}}\right] = \log\left[\frac{R_{t+1}^{(k),*}}{R_t^{f,*}}\frac{R_t^{f,*}}{R_t^f}\frac{S_t}{S_{t+1}}\right] = \log\left[\frac{R_{t+1}^{(k),*}}{R_t^{f,*}}\right] + \log\left[\frac{R_t^{f,*}}{R_t^f}\frac{S_t}{S_{t+1}}\right] = rx_{t+1}^{(k),*} + rx_{t+1}^{FX}. \quad (3)$$

Taking conditional expectations, the total term premium in dollars consists of a foreign bond risk premium, $E_t[rx_{t+1}^{(k),*}]$, plus a currency risk premium, $E[rx_{t+1}^{FX}] = r_t^{f,*} - r_t^f - E_t[\Delta s_{t+1}]$.

2.3 Time-Series Predictability of Foreign Bond Returns

To study the properties of the cross-country differences in expected bond excess returns, we first run individual currency predictability regressions on variables that can be used to predict bond and currency returns. As noted by Ang and Chen (2010) and Berge, Jordà, and Taylor (2011), the intersection of currency and bond return predictors reduces to the level and slope of the term structure, so we focus on those two predictive variables.

We regress the 10-year dollar bond log excess return differential ($rx_{t+1}^{(10),\$} - rx_{t+1}^{(10)}$) on the short-term interest rate differential ($r_t^{f,*} - r_t^f$, Panel A of Table 1) and on the yield curve slope differential ($[y_t^{(10,*)} - y_t^{(1,*)}] - [y_t^{(10)} - y_t^{(1)}]$, Panel B of Table 1), focusing on the post-Bretton Woods sample (1975 – 2015). Given that the 10-year dollar bond log excess return differential (left columns) can be decomposed into the sum of currency log excess returns (rx_{t+1}^{FX}) and local currency bond log excess return differentials ($rx_{t+1}^{(10),*} - rx_{t+1}^{(10)}$), we also regress each of those two components on the same predictors (middle and right columns, respectively). By construction, the sum of the slope coefficients in the middle and right columns equals the slope coefficients in the corresponding left column. For each individual country regression, we report Newey-West standard errors with two lags. The panel regressions include country fixed effects and standard errors are calculated using the Driscoll and Kraay (1998) methodology, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation.

When using interest rate differentials as predictors, there is no consistent evidence in support of predictability of 10-year bond return differentials in dollars: indeed, out of nine countries, there is evidence of return predictability only for Japan. To shed light into the lack of predictability, we consider the decomposition into currency log returns and local currency bond log excess return differentials. As seen in the table, currency log excess returns are strongly forecastable by interest rate differentials: as documented in the existing literature, higher than usual interest rate differentials in a given country pair predict higher than usual currency log excess returns. But, while Treasury bill return differentials in U.S. dollars are forecastable, long-term bond return differentials in U.S. dollars are not. The deterioration of return predictability for long-maturity bonds, compared to Treasury bills, appears to be due to the offsetting effect of local currency bond returns: higher than usual interest rate differences in a given country predict lower local currency bond return differences. In the panel regression, the local bond return slope coefficient is -1.34, largely offsetting the 1.98 slope coefficient in the currency excess return regression. The net effect on dollar bond returns is only 65 basis points, the slope coefficient is not statistically significant, and the regression R^2 is only 0.14%. Thus, from the perspective of a U.S. investor, the time variation in the currency excess return is largely offset by the variation in the local term premium. Note that the local currency log bond excess return differential is highly predictable by the interest rate differential (right column), simply because the interest rate spread in effect predicts itself, as it is a component (with a negative sign) of the dependent variable. When we regress the local currency log return differential on the interest rate differential (not reported), there is no evidence of predictability.

When using yield curve slope differentials as predictors, a similar finding emerges. On the one hand, currency log excess returns are forecastable by yield curve slope differentials: a steeper than usual slope in a given country predicts lower than usual currency log excess returns. On the other hand, a slope steeper than usual in a given country also predicts higher local currency bond excess returns. In the panel regression, the local bond excess return slope coefficient is 3.96, more than offsetting the -2.02 slope coefficient in the currency excess return regression. The net effect on dollar bond excess return differences is 194 basis points in a surprising direction: a steeper slope seems to weakly forecast higher dollar returns for foreign bonds, rather than lower dollar returns as for foreign Treasury Bills. The slope coefficient is statistically significant, but the regression R^2 is only 0.60%. Thus, from the perspective of a U.S. investor, the time variation in the currency excess return is more than offset by the variation in the local term premium. Again, the predictability of local currency bond log excess return differentials (right column) disappears when we attempt to predict local currency log return differentials (not reported).

2.4 Cross-Sectional Properties of Foreign Bond Returns

After focusing on time-series predictability, we turn now to cross-sectional evidence. We sort countries into three portfolios on the level of the short-term interest rates or the slope of their yield curves. Portfolios are

Table 1: Time-Series Predictability

	Bond dollar return difference $rx^{(10),\$} - rx^{(10)}$			Currency excess return rx^{FX}			Bond local currency return diff. $rx^{(10),*} - rx^{(10)}$			Obs.
	β	s.e.	R^2(%)	β	s.e.	R^2(%)	β	s.e.	R^2(%)	
				Panel A: Short-Term Interest Rates						
Australia	-0.15	[1.04]	-0.20	1.29*	[0.66]	0.56	-1.44**	[0.63]	1.51	492
Canada	-1.10	[0.71]	0.11	1.22**	[0.53]	0.46	-2.32***	[0.45]	3.64	492
Germany	1.52	[1.16]	0.37	2.49**	[0.98]	1.71	-0.97	[0.72]	0.48	492
Japan	2.37***	[0.83]	1.13	3.11***	[0.67]	3.48	-0.74	[0.54]	0.13	492
New Zealand	0.69	[0.88]	-0.03	2.23***	[0.53]	3.14	-1.54**	[0.64]	1.62	492
Norway	0.72	[0.66]	0.08	1.74***	[0.59]	2.26	-1.02**	[0.49]	0.97	492
Sweden	-0.64	[0.86]	-0.02	0.89	[0.89]	0.25	-1.53***	[0.53]	2.02	492
Switzerland	1.16	[0.79]	0.33	2.45***	[0.78]	2.43	-1.29***	[0.47]	1.69	492
United Kingdom	1.02	[1.21]	0.04	2.69***	[0.91]	2.44	-1.67**	[0.74]	1.39	492
Panel	0.65	[0.49]	0.14	1.98***	[0.45]	1.96	-1.34***	[0.33]	1.47	4428
				Panel B: Yield Curve Slopes						
Australia	3.84**	[1.57]	1.54	-1.00	[1.18]	-0.02	4.84***	[0.91]	7.65	492
Canada	4.04***	[0.99]	2.25	-0.72	[0.69]	-0.07	4.76***	[0.60]	9.09	492
Germany	0.50	[1.68]	-0.18	-3.05**	[1.38]	1.15	3.55***	[1.10]	4.07	492
Japan	-0.32	[1.32]	-0.19	-4.18***	[1.05]	2.91	3.85***	[0.83]	3.96	492
New Zealand	2.94	[1.89]	1.26	-1.60	[1.08]	0.62	4.55***	[1.16]	7.41	492
Norway	0.59	[1.02]	-0.12	-2.03**	[0.91]	1.33	2.62***	[0.66]	3.35	492
Sweden	3.12***	[1.17]	2.12	-0.13	[1.14]	-0.20	3.25***	[0.77]	5.29	492
Switzerland	0.97	[1.17]	-0.06	-3.59***	[1.29]	1.97	4.55***	[0.76]	8.82	492
United Kingdom	1.59	[1.52]	0.17	-3.17**	[1.25]	2.11	4.75***	[0.88]	7.95	492
Panel	1.94**	[0.79]	0.60	-2.02***	[0.70]	0.98	3.96***	[0.50]	6.17	4428

Notes: The table reports regression results of the bond dollar return difference ($rx_{t+1}^{(10),\$} - rx_{t+1}^{(10)}$, left panel) or the currency excess return (rx_{t+1}^{FX}, middle panel) or the bond local currency return difference ($rx_{t+1}^{(10),*} - rx_{t+1}^{(10)}$, right panel) on the difference between the foreign nominal interest rate and the U.S. nominal interest rate ($r_t^{f,*} - r_t^f$, Panel A) or difference between the foreign nominal yield curve slope and the U.S. nominal yield curve slope ($[y_t^{(10,*)} - y_t^{(1,*)}] - [y_t^{(10)} - y_t^{(1)}]$, Panel B). We use returns on 10-year coupon bonds. The holding period is one month and returns are sampled monthly. The log returns and the yield curve slope differentials are annualized. The sample period is 1/1975–12/2015. The balanced panel consists of Australia, Canada, Japan, Germany, Norway, New Zealand, Sweden, Switzerland, and the U.K. In individual country regressions, standard errors are obtained with a Newey-West approximation of the spectral density matrix with two lags. Panel regressions include country fixed effects, and standard errors are obtained using the Driscoll and Kraay (1998) methodology with two lags. One, two, and three stars denote statistical significance at the 10%, 5%, and 1% level, respectively.

rebalanced every month and those formed at date t only use information available at that date. Portfolio-level log excess returns are obtained by averaging country-level log excess returns across all countries in the portfolio. We first describe results obtained with the 10-year bond indices, reported in Table 2, and then turn to the zero-coupon bonds to study the whole term structure, presented in Figure 1.

2.4.1 Sorting by Interest Rates

We start with the currency portfolios sorted by short-term interest rates. In order to focus on the conditional carry trade, we use interest rates in deviation from their past 10-year rolling mean as the sorting variable. Thus, the first (third) portfolio includes the conditionally low (high) interest rate currencies. Over the 1/1975–12/2015 period, the countries in the third portfolio tend to experience higher inflation rates and lower yield slopes (Panel A of Table 2, left section). Clearly, the classic uncovered interest rate parity condition fails in the cross-section: the currencies in the third portfolio only depreciate by 60 basis points per year on average, not enough to offset the 2.65% interest rate difference. As a result, average currency excess returns increase from low- to high-interest-rate portfolios, ranging from −0.61% to 2.04% per year over the last 40 years. Thus, the long-short currency carry trade (invest in Portfolio 3, short Portfolio 1) implemented with Treasury bills delivers an average annual log excess return of 2.65% and a Sharpe ratio of 0.36, higher than the Sharpe ratio on the U.S. S&P500 equity index over the same period (Panel B). However, implementing the carry trade with long-term bonds does not yield a similar performance, as local currency bond premia decrease from low- to high-interest-rate portfolios, from 3.53% to −0.25%, implying that the 10-year bond carry trade entails a local currency bond premium of −3.78% (Panel C). Thus, the average log excess return on foreign bonds in U.S. dollars of the high-interest-rate portfolio is 1.12% lower than the average excess return of the low-interest-rate portfolio (Panel D); that dollar spread comprises the 2.65% currency risk premium spread and the −3.78% local term premium spread. Therefore, the long-short currency carry trade implemented with long-term government bonds delivers a negative average return, in contrast to the equivalent trade using Treasury bills. Notably, the carry trade dollar return of −1.12% is not statistically significant, suggesting that investors have no reason to favor the long-term bonds of a particular set of countries on the basis of average returns after converting the returns into the same currency (here, U.S. dollars).

Table 2: Cross-sectional Predictability: Bond Portfolios

Portfolio		Sorted by Short-Term Interest Rates 1	2	3	$3-1$	Sorted by Yield Curve Slopes 1	2	3	$1-3$
		Panel A: Portfolio Characteristics							
Inflation rate	Mean	2.90	3.45	4.81	1.91	4.89	3.41	2.87	2.02
	s.e.	[0.16]	[0.19]	[0.23]	[0.20]	[0.23]	[0.19]	[0.18]	[0.19]
Rating	Mean	1.45	1.25	1.49	0.04	1.54	1.38	1.28	0.25
	s.e.	[0.02]	[0.02]	[0.02]	[0.04]	[0.02]	[0.02]	[0.02]	[0.03]
Rating (adj. for outlook)	Mean	1.50	1.37	1.84	0.33	1.84	1.50	1.37	0.47
	s.e.	[0.03]	[0.02]	[0.02]	[0.04]	[0.02]	[0.02]	[0.02]	[0.03]
$y_t^{(10),*} - r_t^{*,f}$	Mean	1.52	0.92	-0.44	-1.96	-0.81	0.85	1.96	-2.76
		Panel B: Currency Excess Returns							
$-\Delta s_{t+1}$	Mean	-0.44	0.11	-0.60	-0.16	-0.95	0.38	-0.36	-0.58
$r_t^{f,*} - r_t^f$	Mean	-0.17	0.54	2.65	2.81	3.35	0.55	-0.88	4.23
rx_{t+1}^{FX}	Mean	-0.61	0.66	2.04	2.65	2.41	0.92	-1.24	3.65
	s.e.	[1.35]	[1.44]	[1.36]	[1.14]	[1.48]	[1.38]	[1.40]	[1.18]
	SR	-0.07	0.07	0.23	0.36	0.26	0.11	-0.14	0.49
		Panel C: Local Currency Bond Excess Returns							
$rx_{t+1}^{(10),*}$	Mean	3.53	2.60	-0.25	-3.78	-1.01	2.29	4.61	-5.61
	s.e.	[0.69]	[0.69]	[0.73]	[0.77]	[0.76]	[0.69]	[0.70]	[0.74]
	SR	0.80	0.58	-0.05	-0.77	-0.21	0.53	1.00	-1.18
		Panel D: Dollar Bond Excess Returns							
$rx_{t+1}^{(10),\$}$	Mean	2.92	3.26	1.80	-1.12	1.40	3.21	3.36	-1.96
	s.e.	[1.56]	[1.58]	[1.57]	[1.33]	[1.64]	[1.57]	[1.62]	[1.38]
	SR	0.29	0.32	0.18	-0.13	0.14	0.33	0.32	-0.22
$rx_{t+1}^{(10),\$} - rx_{t+1}^{(10)}$	Mean	0.14	0.48	-0.98	-1.12	-1.38	0.43	0.59	-1.96
	s.e.	[1.64]	[1.64]	[1.73]	[1.33]	[1.81]	[1.63]	[1.75]	[1.38]

Notes: The countries are sorted by the level of their short term interest rates in deviation from the 10-year mean into three portfolios (left section) or the slope of their yield curves (right section). The slope of the yield curve is measured by the difference between the 10-year yield and the one-month interest rate. The standard errors (denoted s.e. and reported between brackets) were generated by bootstrapping 10,000 samples of non-overlapping returns. The table reports the average change in exchange rates (Δs), the average interest rate difference ($r^{f,*} - r^f$), the average slope of the yield curve ($y^{(10),*} - r^{*,f}$), the average inflation rate, the average currency excess return (rx^{FX}), the average foreign bond excess return on 10-year government bond indices in foreign currency ($rx^{(10),*}$) and in U.S. dollars ($rx^{(10),\$}$), as well as the difference between the average foreign bond excess return in U.S. dollars and the average U.S. bond excess return ($rx^{(10),\$} - rx^{(10)}$). For the excess returns, the table also reports their Sharpe ratios (denoted SR). The holding period is one month. The log returns are annualized. The balanced panel consists of Australia, Canada, Japan, Germany, Norway, New Zealand, Sweden, Switzerland, and the U.K. The data are monthly and the sample is 1/1975–12/2015.

2.4.2 Sorting by Slopes

Similar results emerge when we sort countries into portfolios by the slope of their yield curves. There is substantial turnover in these portfolios, more so than in the usual interest rate-sorted portfolios. On average, the flat slope currencies (first portfolio) tend to be high interest rate currencies, while the steep slope currencies (third portfolio) tend to be low interest rate currencies. As expected, average currency log excess returns decline from 2.41% per annum on the first portfolio to −1.24% per annum on the third portfolio over the last 40 years (Panel B). Therefore, a long-short position of investing in flat-yield-curve currencies (Portfolio 1) and shorting steep-yield-curve currencies (Portfolio 3) delivers a currency excess return of 3.65% per annum and a Sharpe ratio of 0.49. Our findings confirm those of Ang and Chen (2010): the slope of the yield curve predicts currency excess returns at the short end of the maturity spectrum. However, those currency premia are offset by term premia, as local currency bond excess returns and currency excess returns move in opposite directions across portfolios. In particular, the first portfolio produces negative bond average excess returns of −1.01% per year, compared to 4.61% on the third portfolio (Panel C), so the slope carry trade generates an average local currency bond excess return of −5.61% per year. This result is not mechanical: the spread in the slopes is about half of the spread in local currency excess returns. The corresponding average dollar bond excess returns range from 1.40% to 3.36%, so the slope carry trade implemented with 10-year bonds delivers an average annual dollar excess return of −1.96% (Panel D), which is not statistically significant.

2.4.3 Looking Across Maturities

The results we discussed previously focus on the 10-year maturity. We now turn to the full maturity spectrum, using the zero-coupon bond dataset. Our findings are presented graphically in Figure 1, which shows the dollar log excess returns as a function of the bond maturity, using the same set of funding and investment currencies. Investing in the short-maturity bills of countries with flat yield curves (mostly high short-term interest rate countries), while borrowing at the same horizon in countries with steep yield curves (mostly low short-term interest rate countries) leads to an average dollar excess return of 2.09% per year. This is the slope version of the standard currency carry trade. However, when we implement the same strategy using longer maturity bonds instead of short-term bills, the dollar excess return decreases monotonically in bond maturity. At the long end (15-year maturity), the bond term premium more than offsets the currency premium, so the slope carry trade yields a (non-significant) average annual dollar return of −2.18%. Therefore, carry trade strategies that yield positive average excess returns when implemented with short-maturity bonds yield lower (or even negative) excess returns when implemented using long-maturity bonds.

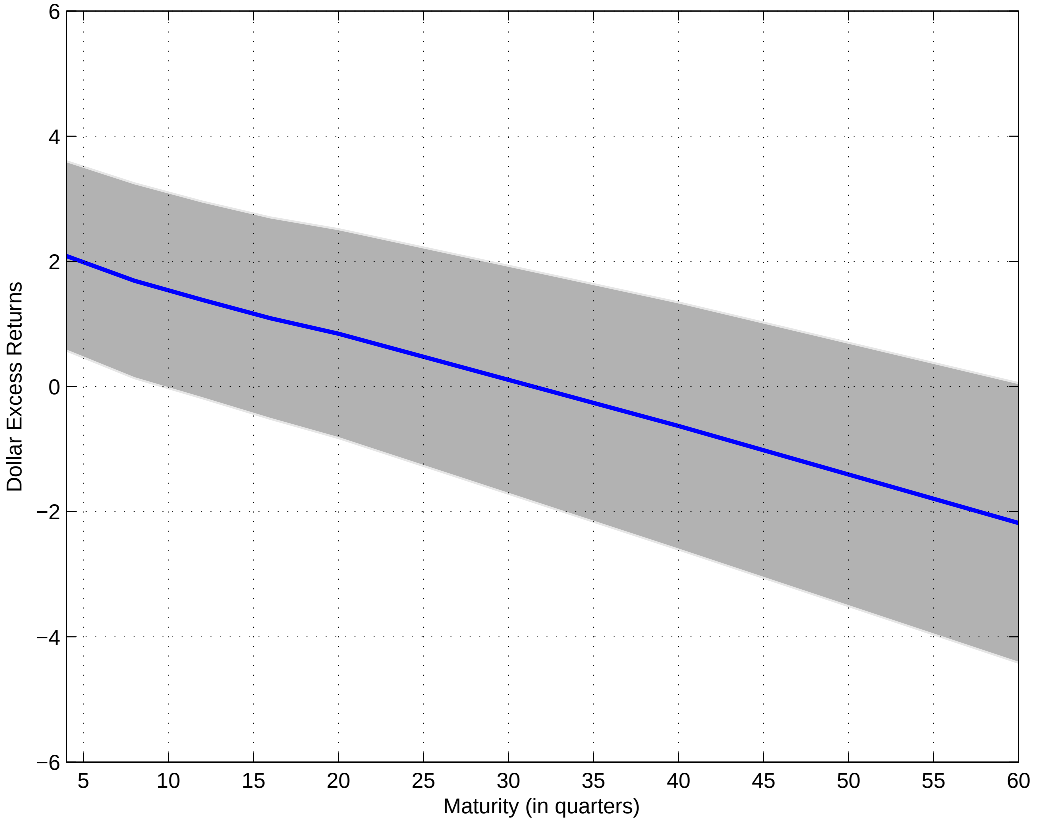

Figure 1: Long-Minus-Short Foreign Bond Risk Premia in U.S. Dollars— The figure shows the dollar log excess returns as a function of the bond maturities. Dollar excess returns correspond to the holding period returns expressed in U.S. dollars of investment strategies that go long and short foreign bonds of different countries. The unbalanced panel of countries consists of Australia, Canada, Japan, Germany, Norway, New Zealand, Sweden, Switzerland, and the U.K. At each date t, the countries are sorted by the slope of their yield curves into three portfolios. The first portfolio contains countries with flat yield curves (mostly high interest rate countries) while the last portfolio contains countries with steep yield curves (mostly low interest rate countries). The first portfolio correspond to the investment currencies while the third one corresponds to the funding currencies. The slope of the yield curve is measured by the difference between the 10-year yield and the 3-month interest rate at date t. The holding period is one quarter. The returns are annualized. The shaded areas correspond to one standard deviation above and below each point estimate. Standard deviations are obtained by bootstrapping 10,000 samples of non-overlapping returns. Zero-coupon data are monthly, and the sample window is 4/1985–12/2015.

2.5 Robustness Checks

We consider many robustness checks, studying (i) different lengths of the bond holding period, (ii) different time windows, (iii) different samples of countries, (iv) sorts on non-demeaned interest rate levels, and (v) other potential explanations of excess returns. All the results are reported in the Online Appendix. Here, we simply describe the main findings.

First, our results appear robust to the choice of the bond holding period. Apart from one-month holding periods, as reported in Table 2, we also consider holding periods of three and twelve months. For those longer holding periods, the return patterns are similar and the dollar bond term premia are not statistically significant.

Second, our findings appear robust across time windows. In particular, we consider a longer sample, 1/1951-12/2015, and a shorter sample, 7/1989–12/2015, and find that our main result remains: an investor in short-term Treasury bills enjoys positive returns, while an investor in long-term bonds of the same countries suffers negative returns. Again, currency and local term premia offset each other, and thus average carry trade returns are different at the short end and the long end of the term structure.

Third, our results appear robust across several samples of countries.[4] For each dataset, we report and comment in the Online Appendix detailed results over the full time window and the post-Bretton-Woods sample. The results are very similar to those in our benchmark sample. Introducing more countries adds power to the experiment, but forces us to consider less liquid and more default-prone bond markets.

Fourth, instead of sorting countries into portfolios on their demeaned nominal interest rates, we sort on their actual nominal interest rates. The carry trade risk premium is somewhat larger when sorting on interest rates, as documented by Lustig, Roussanov, and Verdelhan (2011) and Hassan and Mano (2014), as these sorts capture both the conditional carry trade premium (long in currencies with currently high interest rates) and the unconditional carry trade premium (long in currencies with high average interest rates). Since our paper is about the conditional currency carry trade, we focus on the demeaned sorts.

Fifth, it may be the case that the empirical cross-sectional patterns in bond term premia we uncover are driven by differences in exposure to inflation risk or credit risk: portfolios may have high term premia not

[4] With coupon bonds, we consider two additional sets of countries: first, a larger sample of 20 developed countries (Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, and the U.K.), and second, a large sample of 30 developed and emerging countries (the same as above, plus India, Mexico, Malaysia, the Netherlands, Pakistan, the Philippines, Poland, South Africa, Singapore, Taiwan, and Thailand). We also construct an extended version of the zero-coupon dataset which, in addition to the countries of the benchmark sample, includes the following countries: Austria, Belgium, the Czech Republic, Denmark, Finland, France, Hungary, Indonesia, Ireland, Italy, Malaysia, Mexico, the Netherlands, Poland, Portugal, Singapore, South Africa, and Spain. The data for the aforementioned extra countries are sourced from Bloomberg. The starting dates for the additional countries are as follows: 12/1994 for Austria, Belgium, Denmark, Finland, France, Ireland, Italy, the Netherlands, Portugal, Singapore, and Spain, 12/2000 for the Czech Republic, 3/2001 for Hungary, 5/2003 for Indonesia, 9/2001 for Malaysia, 8/2003 for Mexico, 12/2000 for Poland, and 1/1995 for South Africa.

because the bonds are subject to higher interest rate risk, but because they contain the currencies of countries with high inflation risk or high likelihood of sovereign debt default. To determine if that is a valid concern, we calculate average inflation rates and average credit ratings for each portfolio. We report average inflation rates and credit ratings in Table 2 and in the Online Appendix. For both the long sample (1/1951 – 12/2015) and the post-Bretton Woods sample (1/1975 – 12/2015), there is a negative correlation between average term premia and average inflation rates: term premia are higher in low inflation countries. Thus, assuming that there is a positive association between average inflation rates and exposure to inflation risk, inflation risk does not account for our findings. This is true not only for our benchmark set of countries, but also for the extended sets of countries. Similarly, the cross-sectional patterns in term premia we observe empirically are not likely to be due to sovereign default risk. As seen in Table 2, countries with high average local currency bond premia have average credit ratings (both unadjusted and adjusted for outlook) that are either lower than, or very similar to, the ratings of countries with low average local currency bond premia. That finding is robust to considering different sample periods: it holds both in the long sample period (1/1951 – 12/2015) and in the 7/1989–12/2015 period, during which full ratings are available. Therefore, while it is difficult to definitively rule out an inflation- or credit-based risk premium explanation, a simple interest rate risk interpretation seems the most relevant, especially over the last thirty years.

3 The Term Structure of Currency Carry Trade Risk Premia: A Challenge

In this section, we show that the downward sloping term structure of currency carry trade risk premia is a challenge for existing models. Several two-country models replicate the failure of the U.I.P. condition, but they cannot replicate the portfolio evidence on carry trade risk premia. In similar models with multiple countries, investors can diversify away the country-specific exchange rate risk and there are no cross-sectional differences in carry trade returns across portfolios. To the best of our knowledge, only two models can so far replicate the portfolio evidence on carry trades: the multi-country long-run risk model of Colacito, Croce, Gavazzoni, and Ready (2017) and the multi-country reduced-form factor model of Lustig, Roussanov, and Verdelhan (2011). We focus on the latter because of its flexibility and close forms, and revisit the long-run risk model in Section 6, along with other explanations of the U.I.P. puzzle.

3.1 The Necessary Condition for Replicating the U.I.P Puzzle

We start with a review of a key necessary condition on the stochastic discount factors, established by Bekaert (1996) and Bansal (1997) and generalized by Backus, Foresi, and Telmer (2001). To do so, we first introduce some additional notation.

Pricing Kernels and Stochastic Discount Factors The nominal pricing kernel is denoted by $\Lambda_t(\varpi)$; it corresponds to the marginal value of a a currency unit delivered at time t in the state of the world ϖ. The nominal SDF M is the growth rate of the pricing kernel: $M_{t+1} = \Lambda_{t+1}/\Lambda_t$. Therefore, the price of a zero-coupon bond that promises one currency unit k periods into the future is given by:

$$P_t^{(k)} = E_t\left(\frac{\Lambda_{t+k}}{\Lambda_t}\right). \tag{4}$$

SDF Entropy SDFs are volatile, but not necessarily normally distributed. In order to measure the time-variation in their volatility, it is convenient to use entropy, rather than variance (Backus, Chernov, and Zin, 2014). The conditional entropy L_t of any random variable X_{t+1} is defined as:

$$L_t\left(X_{t+1}\right) = \log E_t\left(X_{t+1}\right) - E_t\left(\log X_{t+1}\right). \tag{5}$$

If X_{t+1} is conditionally lognormally distributed, then the conditional entropy is equal to one half of the conditional variance of the log of X_{t+1}: $L_t\left(X_{t+1}\right) = (1/2)var_t\left(\log X_{t+1}\right)$. If X_{t+1} is not conditionally lognormal, the entropy also depends on the higher moments: $L_t(X_{t+1}) = \kappa_{2t}/2! + \kappa_{3t}/3! + \kappa_{4t}/4! + \ldots$, where $\{\kappa_{it}\}_{i=2}^{\infty}$ are the cumulants of $\log X_{t+1}$.

Exchange Rates When markets are complete, the change in the nominal exchange rate corresponds to the ratio of the domestic to foreign nominal SDFs:

$$\frac{S_{t+1}}{S_t} = \frac{\Lambda_{t+1}}{\Lambda_t}\frac{\Lambda_t^*}{\Lambda_{t+1}^*}. \tag{6}$$

The no-arbitrage definition of the exchange rate comes directly from the Euler equations of the domestic and foreign investors, for any asset return R^* expressed in foreign currency terms: $E_t[M_{t+1}R_{t+1}^* S_t/S_{t+1}] = 1$ and

$E_t[M^*_{t+1}R^*_{t+1}] = 1$. When markets are complete, the SDF is unique, and thus the change in the exchange rate equals the ratio of the two SDFs. When markets are incomplete, there are other candidate exchange rates.

Currency Risk Premia As Bekaert (1996) and Bansal (1997) show, in models with lognormally distributed SDFs the conditional log currency risk premium $E_t(rx^{FX})$ equals the half difference between the conditional variance of the log domestic and foreign SDFs. This result can be generalized to non-Gaussian economies.[5] When higher moments matter and markets are complete, the currency risk premium is equal to the difference between the conditional entropy of two SDFs (Backus, Foresi, and Telmer, 2001):

$$E_t\left(rx^{FX}_{t+1}\right) = r^{f,*}_t - r^f_t - E_t(\Delta s_{t+1}) = L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) - L_t\left(\frac{\Lambda^*_{t+1}}{\Lambda^*_t}\right). \tag{7}$$

According to the U.I.P. condition, expected changes in exchange rates should be equal to the difference between the home and foreign interest rates and, thus, the currency risk premium should be zero. In the data, the currency risk premium is as large as the equity risk premium. Any no arbitrage model that addresses the U.I.P. puzzle must thus satisfy a simple necessary condition: high interest rate countries must exhibit relatively less volatile SDFs. In the absence of differences in conditional volatility, no arbitrage models are unable to generate a currency risk premium and the U.I.P. counterfactually holds in the model economy.

3.2 An Example: A Reduced-Form Factor Model

We now turn to a flexible N-country, reduced-form model that can both replicate the deviations from U.I.P. and generate large currency carry trade returns on currency portfolios. To replicate the portfolio evidence, as Lustig, Roussanov, and Verdelhan (2011) show, no arbitrage models need to incorporate global shocks to the SDFs along with country heterogeneity in the exposure to those shocks. Following Lustig, Roussanov, and Verdelhan (2014), we consider a world with N countries and currencies in a setup inspired by classic term structure

[5] The literature on disaster risk in currency markets shows that higher order moments are critical for understanding currency returns. In earlier work, Brunnermeier, Nagel, and Pedersen (2009) show that risk reversals increase with interest rates. Jurek (2014) provides a comprehensive empirical investigation of hedged carry trade strategies. Gourio, Siemer, and Verdelhan (2013) study a real business cycle model with disaster risk. Farhi, Fraiberger, Gabaix, Ranciere, and Verdelhan (2013) estimate a no-arbitrage model with crash risk using a cross-section of currency options. Chernov, Graveline, and Zviadadze (2011) study jump risk at high frequencies. Gavazzoni, Sambalaibat, and Telmer (2012) show that lognormal models cannot account for the cross-country differences in carry returns and interest rate volatilities.

models.[6] In the model, the risk prices associated with country-specific shocks depend only on country-specific factors, but the risk prices of world shocks depend on world and country-specific factors. To describe these risk prices, the authors introduce a common state variable z_t^w, shared by all countries, and a country-specific state variable z_t^i. The country-specific and world state variables follow autoregressive square-root processes:

$$\begin{aligned} z_{t+1}^i &= (1-\phi)\theta + \phi z_t^i - \sigma\sqrt{z_t^i}u_{t+1}^i, \\ z_{t+1}^w &= (1-\phi^w)\theta^w + \phi^w z_t^w - \sigma^w\sqrt{z_t^w}u_{t+1}^w. \end{aligned}$$

Lustig, Roussanov, and Verdelhan (2014) assume that in each country i, the logarithm of the real SDF $\widetilde{m}^i$ follows a three-factor conditionally Gaussian process:

$$-\widetilde{m}_{t+1}^i = \alpha + \chi z_t^i + \sqrt{\gamma z_t^i}u_{t+1}^i + \tau z_t^w + \sqrt{\delta^i z_t^w}u_{t+1}^w + \sqrt{\kappa z_t^i}u_{t+1}^g,$$

where u_{t+1}^i is a country-specific SDF shock, while u_{t+1}^w and u_{t+1}^g are common to all countries' SDFs. All three innovations are i.i.d. Gaussian, with zero mean and unit variance. To be parsimonious, Lustig, Roussanov, and Verdelhan (2014) limit the heterogeneity in the SDF parameters to the different loadings δ^i on the world shock u_{t+1}^w; all the other parameters are identical for all countries. Therefore, the model is a restricted version of the multi-factor dynamic term structure models, and there exist closed form solutions for bond yields and risk premia.

There are two types of common shocks. The first type, u_{t+1}^w, is priced proportionally to country exposure δ^i, and since δ^i is a fixed characteristic of country i, differences in such exposure are *permanent*. The second type, u_{t+1}^g, is priced proportionally to z_t^i, so heterogeneity with respect to this innovation is *transitory*: all countries are equally exposed to this shock *on average*, but conditional exposures vary over time and depend on country-specific economic conditions. Finally, the real risk-free rate is $\widetilde{r}_t^{f,i} = \alpha + \left(\chi - \frac{1}{2}(\gamma+\kappa)\right) z_t^i + \left(\tau - \frac{1}{2}\delta^i\right) z_t^w$.

Country i's inflation process is given by $\pi_{t+1}^i = \pi_0 + \eta^w z_t^w + \sigma_\pi \epsilon_{t+1}^i$, where the inflation innovations ϵ_{t+1}^i are i.i.d. Gaussian. It follows that the log nominal risk-free rate in country i is given by $r_t^{f,i} = \pi_0 + \alpha + \left(\chi - \frac{1}{2}(\gamma+\kappa)\right) z_t^i + \left(\tau + \eta^w - \frac{1}{2}\delta^i\right) z_t^w - \frac{1}{2}\sigma_\pi^2$. The nominal bond prices in logs are affine in the state variable z

[6] In the Online Appendix, we cover a wide range of term structure models, from the seminal Vasicek (1977) model to the classic Cox, Ingersoll, and Ross (1985) model and to the most recent, multi-factor dynamic term structure models. To save space, we focus here on their most recent international finance version, illustrated in Lustig, Roussanov, and Verdelhan (2014).

and z^w: $p_t^{(n),i} = -C_0^{n,\$,i} - C_1^{n,\$} z_t - C_2^{n,\$,i} z_t^w$, where the loadings $(C_0^{n,\$,i}, C_1^{n,\$}, C_0^{n,\$,i})$ are defined in the Appendix. Condition (7) implies that the foreign currency risk premium is given by:

$$E_t(rx_{t+1}^{FX,i}) = -\frac{1}{2}(\gamma + \kappa)(z_t^i - z_t) + \frac{1}{2}(\delta - \delta^i) z_t^w.$$

Investors obtain high foreign currency risk premia when investing in currencies with relative small exposure to the two global shocks. That is the source of short-term carry trade risk premia.

3.3 The Model-Implied Term Structure of Currency Carry Trade Risk Premia

Using the benchmark calibration of the Lustig, Roussanov, and Verdelhan (2014) model, we calculate the model-implied term structure of currency risk premia when implementing the slope carry trade strategy (invest in low yield slope currencies, short the high yield slope interest rate currencies). This is very similar to investing in high interest rate countries while borrowing in low interest rate countries. The simulation details are provided in the Appendix.

Figure 2, obtained with simulated data, is the model counterpart to Figure 1, obtained with actual data. A clear message emerges: while this model produces U.I.P. deviations (and thus currency risk premia) at the short end of the yield curve, the model produces a flat term structure of currency carry trade risk premia. We turn now to a novel necessary condition that future models need to satisfy in order to generate a downward-sloping term structure.

4 Foreign Long-Term Bond Returns and the Properties of SDFs

In this section, we derive a novel, preference-free necessary condition that no-arbitrage models need to satisfy in order to reproduce the downward sloping term structure of currency carry trade risk premia. To do so, we first review a useful decomposition of the pricing kernel into a permanent and a transitory component.

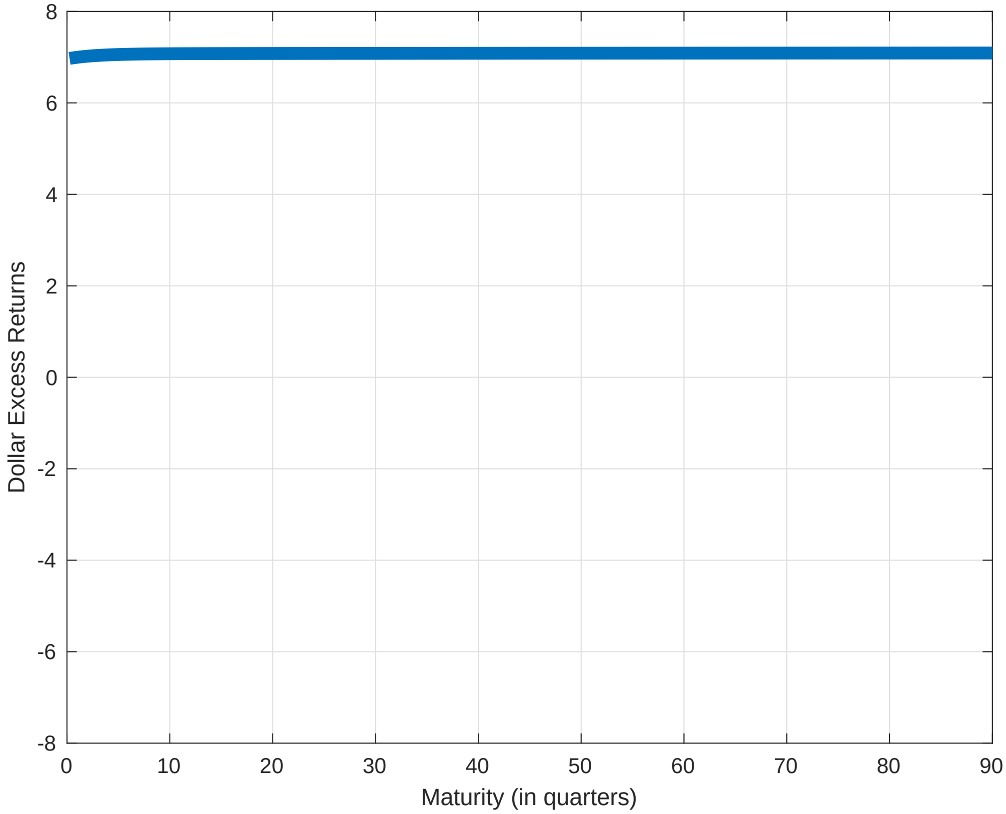

Figure 2: Simulated Long-Minus-Short Foreign Bond Risk Premia in U.S. Dollars— The figure shows the simulated average dollar log excess return of the slope carry trade strategy as a function of the bond maturities in the reduced-form model of Lustig, Roussanov, and Verdelhan (2014). At each date t, currencies are sorted into three portfolios by the slope of their yield curve (measured as the difference between the 10-year and the three-month yields). The first portfolio contains the currencies of countries with low yield slopes, while the third portfolio contains the currencies of countries with high yield slope. The slope carry trade strategy invests in the first portfolio and shorts the third portfolio. The model is simulated at the monthly frequency. The holding period is one month and returns are annualized.

4.1 Pricing Kernel Decomposition

Our results build on the Alvarez and Jermann (2005) decomposition of the pricing kernel Λ_t into a permanent component $\Lambda_t^{\mathbb{P}}$ and a transitory component $\Lambda_t^{\mathbb{T}}$ using the price of the long-term bond:

$$\Lambda_t = \Lambda_t^{\mathbb{P}} \Lambda_t^{\mathbb{T}}, \text{ where } \Lambda_t^{\mathbb{T}} = \lim_{k \to \infty} \frac{\delta^{t+k}}{P_t^{(k)}}, \tag{8}$$

where the constant δ is chosen to satisfy the following regularity condition: $0 < \lim_{k\to\infty} \frac{P_t^{(k)}}{\delta^k} < \infty$ for all t.[7] Alvarez and Jermann (2005) assume that, for each $t+1$, there exists a random variable x_{t+1} with finite expected value $E_t(x_{t+1})$ such that almost surely $\frac{\Lambda_{t+1}}{\delta^{t+1}} \frac{P_{t+1}^{(k)}}{\delta^k} \leq x_{t+1}$ for all k. Under those regularity conditions, the infinite-maturity bond return is:

$$R_{t+1}^{(\infty)} = \lim_{k\to\infty} R_{t+1}^{(k)} = \lim_{k\to\infty} P_{t+1}^{(k-1)}/P_t^{(k)} = \frac{\Lambda_t^{\mathbb{T}}}{\Lambda_{t+1}^{\mathbb{T}}}. \tag{9}$$

The permanent component, $\Lambda_t^{\mathbb{P}}$, is a martingale and is an important part of the pricing kernel: Alvarez and Jermann (2005) derive a lower bound of its volatility and, given the size of the equity premium relative to the term premium, conclude that it accounts for most of the SDF volatility.[8] Thus, a lot of persistence in the pricing kernel is needed to jointly deliver a low term premium and a high equity premium.

4.2 Main Preference-Free Result on Long-Term Bond Returns

We now use this SDF decomposition to understand the properties of the dollar returns of long-term bonds. Recall that the dollar term premium on a foreign bond position, denoted by $E_t[rx_{t+1}^{(k),\$}]$, can be expressed as the sum of foreign term premium in local currency terms, $E_t[rx_{t+1}^{(k),*}]$, plus a currency risk premium, $E_t[rx_{t+1}^{FX}] = r_t^{f,*} - r_t^f - E_t[\Delta s_{t+1}]$. Here, we consider the dollar term premium of an infinite-maturity foreign bond, so we let $k \to \infty$.

Proposition 1. *The foreign term premium on the long-term bond in dollars is equal to the domestic term premium plus the difference between the domestic and foreign entropies of the permanent components of the stochastic discount factors:*

$$E_t[rx_{t+1}^{(\infty),\$}] = E_t\left[rx_{t+1}^{(\infty)}\right] + L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) - L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P},*}}{\Lambda_t^{\mathbb{P},*}}\right). \tag{10}$$

[7] The SDF representation defined here is subject to important limitations that need to be highlighted. Hansen and Scheinkman (2009) point out that this decomposition is not unique under the assumptions used in Alvarez and Jermann (2005). The transitory and permanent components are potentially correlated, which may complicate their interpretation. Despite this limitation, this representation proves to be particularly useful when analyzing short-horizon returns on longer maturity bonds. We follow the more general Hansen and Scheinkman (2009) decomposition when studying bond yields. The two methods lead to identical decompositions for the models that we study in this paper.

[8] Note that $\Lambda_t^{\mathbb{P}}$ is equal to:

$$\Lambda_t^{\mathbb{P}} = \lim_{k\to\infty} \frac{P_t^{(k)}}{\delta^{t+k}} \Lambda_t = \lim_{k\to\infty} \frac{E_t(\Lambda_{t+k})}{\delta^{t+k}}.$$

The second regularity condition ensures that the expression above is a martingale.

In case of an adverse temporary innovation to the foreign pricing kernel, the foreign currency appreciates, so a domestic position is the foreign long-term zero-coupon bond experiences a capital gain. However, this capital gain is exactly offset by the capital loss suffered on that bond as a result of the increase in foreign interest rates. Hence, interest rate exposure completely hedges the temporary component of the currency risk exposure, and the only source of priced currency risks in the foreign bond position are permanent pricing kernel innovations. Therefore, in order to have a non-zero bond risk premium at long maturities, conditional entropy differences in the permanent component of SDFs are required. If the domestic and foreign SDFs have identical conditional entropy, then high local currency term premia are always associated with low currency risk premia and vice-versa, so dollar term premia are identical across currencies.

To intuitively link the long-run properties of SDFs to foreign bond returns and exchange rates, let us consider the simple benchmark of countries represented by a stand-in agents with power utility and *i.i.d.* consumption growth rates. In that case, all SDF shocks are permanent ($\Lambda_t = \Lambda_t^{\mathbb{P}}$ for all t) and the risk-free rate is constant, so bonds of different maturities offer the same returns. Foreign bond investments differ from domestic bond investments only because of the presence of exchange rate risk and, since consumption growth rates are *i.i.d*, exchange rates are stationary in changes but not in levels. Finally, carry trade excess returns are the same at the short end, (see Equation (7)) and at the long end (see Equation (10)) of the yield curve, so the term structure of currency carry trade risk premia is flat.

Proposition 1 pertains to *conditional* risk premia and is, thus, relevant for interpreting our empirical time series predictability results and the average excess returns of currency portfolios sorted by conditioning information (the level of the short-term interest rate or the slope of the yield curve). In fact, we can easily characterize the condition that needs to be satisfied in order for conditional foreign dollar term premia to be identical to conditional domestic term premia and, therefore, for differences in dollar bond excess returns across countries to lack predictability.

Condition 1. *In order for the conditional dollar term premia on long-maturity bonds to be identical across countries, the conditional entropy of the permanent SDF component also has to be identical across countries:* $L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) = L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P},*}}{\Lambda_t^{\mathbb{P},*}}\right)$, *for all* t.

We refer to this condition as long-run risk neutrality. This condition rules out shocks that have permanent effects on the quantity or price of risk, because these shocks would change the conditional entropy of the permanent SDF component, unless those permanent shocks are identical across countries. If this condition fails, portfolios sorted on conditioning variables produce non-zero currency carry trade risk premia at the long end of the term structure, as the conditional dollar term premia of long-maturity bonds differ across countries. There is one obvious way to satisfy this condition: having constant (and identical across countries) entropy of the permanent SDF component. But equity and returns rule out this case (see Alvarez and Jermann, 2005; Hansen and Scheinkman, 2009).

To provide some intuition on the long-run risk neutrality condition, we rely on an example from Alvarez and Jermann (2005), who consider a model with conditionally log-normally distributed pricing kernels driven by both permanent and transitory shocks.

Example 1. *Consider the following pricing kernel (Alvarez and Jermann, 2005):*

$$\begin{aligned} \log \Lambda_{t+1}^{\mathbb{P}} &= -\frac{1}{2}\sigma_P^2 + \log \Lambda_t^{\mathbb{P}} + \varepsilon_{t+1}^P, \\ \log \Lambda_{t+1}^{\mathbb{T}} &= \log \beta^{t+1} + \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t+1-i}^T, \end{aligned}$$

where α is a square summable sequence, and ε^P and ε^T are serially independent and normally distributed random variables with mean zero, variance σ_P^2 and σ_T^2, respectively, and covariance σ_{TP}. A similar decomposition applies to the foreign pricing kernel.

In this economy, Alvarez and Jermann (2005) show that the domestic term premium is given by the following expression: $E_t\left[rx_{t+1}^{(\infty)}\right] = \frac{1}{2}\sigma_T^2 + \sigma_{TP}$. Only transitory risk is priced in the market for long-maturity bonds: when marginal utility is transitorily high, interest rates increase because the representative agent wants to borrow, so long-term bonds suffer a capital loss. On the other hand, permanent shocks to marginal utility do not affect the prices of long-term bonds at all. Similarly, the foreign term premium, in local currency terms, is $E_t\left[rx_{t+1}^{(\infty),*}\right] = \frac{1}{2}\left(\sigma_T^*\right)^2 + \sigma_{TP}^*$.

The currency risk premium is the difference in the two countries' conditional SDF entropy:

$$E_t\left[rx_t^{FX}\right] = r_t^{f,*} - r_t^f - E_t[\Delta s_{t+1}] = \frac{1}{2}\left(\sigma_T^2 + 2\sigma_{TP} + \sigma_P^2\right) - \frac{1}{2}\left((\sigma_T^*)^2 + 2\sigma_{TP}^* + (\sigma_P^*)^2\right).$$

As a result, the foreign term premium in dollars, given by Equation (7), is:

$$E_t\left[rx_{t+1}^{(\infty),\$}\right] = E_t\left[rx_{t+1}^{(\infty),*}\right] + E_t\left[rx_t^{FX}\right] = \frac{1}{2}\sigma_T^2 + \sigma_{TP} + \frac{1}{2}\left(\sigma_P^2 - (\sigma_P^*)^2\right).$$

Provided that $\sigma_P^2 = (\sigma_P^*)^2$, Condition 1 is satisfied, and the foreign term premium in dollars equals the domestic term premium:

$$E_t\left[rx_{t+1}^{(\infty),\$}\right] = \frac{1}{2}\sigma_T^2 + \sigma_{TP} = E_t\left[rx_{t+1}^{(\infty)}\right].$$

5 Holding-Period Returns on Long-term Bonds vs Long-Horizon U.I.P.

Examining the conditional moments of one-period returns on long-maturity bonds, the focus of our paper, is not equivalent to studying the moments of long-maturity bond yields in tests of the long-horizon U.I.P. condition. In this section, we show that U.I.P holds *on average* over long horizons when exchange rates are stationary in levels.

5.1 Finite Long-Horizon U.I.P.

The long-horizon U.I.P. condition states that the expected return over k periods on a foreign bond, once converted into domestic currency, is equal to the expected return on a domestic bond over the same investment horizon. The per period log risk premium on a long position in foreign currency over k periods consists of the yield spread minus the per period expected rate of depreciation over those k periods:

$$E_t[rx_{t\to t+k}^{FX}] = y_t^{(k),*} - y_t^{(k)} - \frac{1}{k}E_t[\Delta s_{t\to t+k}]. \tag{11}$$

The long-horizon U.I.P condition states that this risk premium is zero. As is well-known, this risk premium is the sum of a term premium and future currency risk premia. To see that, start from the definition of the one-period currency risk premium: $E_t\left[\Delta s_{t\to t+1}\right] = r_t^{f,*} - r_t^f - E_t\left[rx_{t+1}^{FX}\right]$. Summing up over k periods leads to:

$$E_t[\Delta s_{t\to t+k}] = E_t\left[\sum_{j=1}^{k}\left(r_{t+j-1}^{f,*} - r_{t+j-1}^{f}\right)\right] - E_t\left[\sum_{j=1}^{k} rx_{t+j}^{FX}\right]. \tag{12}$$

From Equations (11) and (12), it follows that the log currency risk premium over k periods is given by:

$$E_t[rx^{FX}_{t\to t+k}] \;=\; (y^{(k),*}_t - y^{(k)}_t) + \frac{1}{k}\sum_{j=1}^{k} E_t\left(r^f_{t+j-1} - r^{f,*}_{t+j-1}\right) + \frac{1}{k}\sum_{j=1}^{k} E_t(rx^{FX}_{t+j}). \tag{13}$$

The first two terms measure the deviations from the expectations hypothesis over the holding period k, whereas the last term measures the deviations from short-run U.I.P. over the k periods.

We can use a multi-horizon version of Equation (7) to show that the currency risk premium over k periods depends on conditional SDF entropy:

$$E_t[rx^{FX}_{t\to t+k}] = \frac{1}{k}\left[L_t\left(\frac{\Lambda_{t+k}}{\Lambda_t}\right) - L_t\left(\frac{\Lambda^*_{t+k}}{\Lambda^*_t}\right)\right]. \tag{14}$$

The expression above states that only differences in k-period conditional SDF entropy give rise to long-run deviations from U.I.P. Therefore, the risk premium on a multi-period long position in foreign currency depends on how quickly SDF entropy builds up domestically and abroad over the holding period. To develop some intuition, we consider a Gaussian example in the Appendix.

5.2 Infinite Horizon U.I.P.

To link long-horizon U.I.P. and exchange rate dynamics, we use again the decomposition of the pricing kernel proposed by Alvarez and Jermann (2005).

Exchange Rate Decomposition Exchange rate changes can be represented as the product of two components, defined below:

$$\frac{S_{t+1}}{S_t} = \left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\frac{\Lambda^{\mathbb{P},*}_t}{\Lambda^{\mathbb{P},*}_{t+1}}\right)\left(\frac{\Lambda^{\mathbb{T}}_{t+1}}{\Lambda^{\mathbb{T}}_t}\frac{\Lambda^{\mathbb{T},*}_t}{\Lambda^{\mathbb{T},*}_{t+1}}\right) = \frac{S^{\mathbb{P}}_{t+1}}{S^{\mathbb{P}}_t}\frac{S^{\mathbb{T}}_{t+1}}{S^{\mathbb{T}}_t}. \tag{15}$$

Exchange rate changes reflect differences in both the transitory and the permanent component of the two countries' SDFs.[9] If two countries share the same martingale component of the pricing kernel, then the resulting exchange rate is stationary and Condition 1 is trivially satisfied. However, exchange rate stationarity is obviously not necessary for long-run risk neutrality.

[9] Note that $S^{\mathbb{P}}_{t+1}$, the ratio of two martingales, is itself not a martingale in general. However, in the class of affine term structure models, this exchange rate component is indeed a martingale.

Infinite-Horizon U.I.P. We now turn to the infinite-horizon case ($k \to \infty$) to show that long-horizon U.I.P holds on average when exchange rates are stationary in levels. When domestic long-term interest rates increase relative to the corresponding foreign rates, it implies that $\Lambda_{t+1}^{\mathbb{T}}$ is high relative to its foreign counterparts, so the transitory component of the exchange rate $S_{t+1}^{\mathbb{T}}$ appreciates. That follows directly from the definition of the transitory component of exchange rate changes, given by

$$\Delta s_{t+1}^{\mathbb{T}} = \left(\lambda_{t+1}^{\mathbb{T}} - \lambda_t^{\mathbb{T}}\right) - \left(\lambda_{t+1}^{\mathbb{T},*} - \lambda_t^{\mathbb{T},*}\right), \tag{16}$$

where $\lambda_t^{\mathbb{T}} \equiv \log \Lambda_t^{\mathbb{T}} = \lim_{k\to\infty}(t+k)\log\delta + \lim_{k\to\infty} k y_t^{(k)}$. As a result, a currency experiences a temporary appreciation when its long-term interest rates increase more than the foreign ones:

$$\Delta s_{t+1}^{\mathbb{T}} = \log\delta - \log\delta^* + \lim_{k\to\infty} k\left(\Delta y_{t+1}^{(k)} - \Delta y_{t+1}^{(k),*}\right). \tag{17}$$

By backward substitution, it follows that the transitory component of the exchange rate in levels is given by the spread in long-term yields:

$$s_t^{\mathbb{T}} = s_0 + t(\log\delta - \log\delta^*) + \lim_{k\to\infty} k\left(y_t^{(k)} - y_t^{(k),*}\right) - \lim_{k\to\infty} k\left(y_0^{(k)} - y_0^{(k),*}\right). \tag{18}$$

In short, the domestic currency is strong when the long-term rate is high at home relative to abroad. As a result, the domestic currency is expected to depreciate in the future by the size of the spread in long-term rates, as the transitory component of the exchange rate reverts back to its mean. The following proposition provides the result formally.

Proposition 2. *The expected rate of transitory depreciation is always given by the spread in the long-term interest rates:*

$$\lim_{k\to\infty}\frac{1}{k}E_t[\Delta s_{t\to t+k}^{\mathbb{T}}] = -\lim_{k\to\infty}\left(y_t^{(k)} - y_t^{(k),*}\right).$$

Thus, the per period currency risk premium for long holding periods is the permanent component of the rate of appreciation:

$$\lim_{k\to\infty} E_t[rx_{t\to t+k}^{FX}] = \lim_{k\to\infty}\left(y_t^{(k),*} - y_t^{(k)}\right) - \lim_{k\to\infty}\frac{1}{k}E_t[\Delta s_{t\to t+k}] = -\lim_{k\to\infty}\frac{1}{k}E_t[\Delta s_{t\to t+k}^{\mathbb{P}}].$$

It follows that, if exchange rates are stationary in levels, in which case the permanent component of exchange rate changes is zero, then the long-run log currency risk premium converges to zero as $k \to \infty$ and long-run U.I.P. holds. In this case, the slope coefficient in the regression of long-run exchange rate changes on yield differences converges to one and the intercept converges to zero. In other words, if exchange rates are stationary, then long-run U.I.P. holds. This result is previewed in Backus, Boyarchenko, and Chernov (2016), who show that claims to stationary cash flows earn a zero log risk premium over long holding periods. It follows that long-run deviations from U.I.P. are consistent with no arbitrage only if the exchange rate is not stationary in levels.

We can further develop our understanding of the infinite-horizon U.I.P. condition, but we need to limit ourselves to *unconditional* risk premia. Unconditionally, under some additional assumptions, as the maturity of the bonds and the holding period increase, the risk of the permanent SDF component dominates. In the limit ($k \to \infty$), only the differences in the entropy of the martingale SDF components survive. The following proposition states these assumptions and this result precisely.

Proposition 3. *If the stochastic discount factors* $\frac{\Lambda_{t+1}}{\Lambda_t}$ *and* $\frac{\Lambda^*_{t+1}}{\Lambda^*_t}$ *are strictly stationary, and* $\lim_{k\to\infty} \frac{1}{k} L\left(E_t\left[\frac{\Lambda_{t+k}}{\Lambda_t}\right]\right) = 0$ *and* $\lim_{k\to\infty} \frac{1}{k} L\left(E_t\left[\frac{\Lambda^*_{t+k}}{\Lambda^*_t}\right]\right) = 0$*, then the per period long-run currency risk premium is given by:*

$$\lim_{k\to\infty} E[rx^{FX}_{t\to t+k}] = \lim_{k\to\infty} E\left(y^{(k),*}_t - y^{(k)}_t\right) - \lim_{k\to\infty} \frac{1}{k} E[\Delta s_{t\to t+k}] = E\left[L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) - L_t\left(\frac{\Lambda^{\mathbb{P},*}_{t+1}}{\Lambda^{\mathbb{P},*}_t}\right)\right].$$

This immediately implies that the per period currency risk premium converges to zero on average, and therefore long-run U.I.P. holds on average, if long-run risk neutrality, i.e. Condition 1, is satisfied. Notably, unconditional long-run U.I.P. does not necessarily require stationary exchange rates, as long-run risk neutrality is a weaker condition that exchange rate stationarity. To the best of our knowledge, we are the first to demonstrate the connection between exchange rate stationarity, long-maturity bond dollar return parity and long-run U.I.P in a no arbitrage framework.

6 Implications for International Finance Models

Armed with our preference-free results, we now review several models that have been proposed to address the U.I.P. puzzle and study their implications for long-term bond returns in U.S. dollars. We consider two-country versions of the habit, long-run risk, and disaster risk models with country-specific shocks and revisit the multi-country factor model studied in Section 3. All the intermediary steps to determine the two components of the SDFs are reported in the Online Appendix. In the habit model, Condition 1 is trivially satisfied because the quantity of permanent risk is constant. In all the other models, Condition 1 implies some novel parameter restrictions that were not satisfied by the original calibrations of these models, thus explaining why they do not match the empirical evidence highlighted in Section 2.

6.1 External Habit Model

We first consider a version of the Campbell and Cochrane (1999) model with external habit used by Wachter (2006), Verdelhan (2010), and Stathopoulos (2017) to study the properties of interest rates and exchange rates. In this model, the agent maximizes the power utility (with risk aversion coefficient equal to γ and time preference parameter δ) stemming from the difference between consumption and an endogenous subsistence (or habit) level that is a function of past consumption. The log surplus consumption ratio, which measures the percentage distance between consumption and habit, is a stationary variable that evolves as follows: $s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)\varepsilon_{t+1}$, where $\bar{s}$ is its unconditional mean. The sensitivity function λ is given by:

$$\lambda(s_t) = \begin{cases} \frac{1}{\bar{S}}\sqrt{1-2(s_t-\bar{s})} - 1, \text{if } s < s_{max} \\ 0, \text{if } s \geq s_{max} \end{cases},$$

where $\bar{S} = \sigma\sqrt{\frac{\gamma}{1-\phi-B/\gamma}}$ is the steady-state value of the surplus consumption ratio and $s_{max} = \bar{s} + \frac{1}{2}(1-\bar{S}^2)$ is the upper bound of the log surplus consumption ratio. Aggregate consumption growth is given by $\Delta c_{t+1} = g + \varepsilon_{t+1}$, where $\varepsilon_{t+1} \sim N(0, \sigma^2)$. Therefore, the log SDF has law of motion

$$\log\frac{\Lambda_{t+1}}{\Lambda_t} = \log\delta - \gamma g - \gamma(1-\phi)(\bar{s}-s_t) - \gamma(1+\lambda(s_t))\varepsilon_{t+1},$$

and the conditional SDF entropy is

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{\gamma^2\sigma^2}{2}(1+\lambda(s_t))^2 = \frac{\gamma^2\sigma^2}{2}\frac{1}{\bar{S}^2}(1-2(s_t-\bar{s})).$$

The parameter B determines the cyclicality of the real interest rate. In particular, the equilibrium log risk-free rate is given by:

$$r_t^f = -\log\delta + \gamma g - \frac{1}{2}\frac{\gamma^2\sigma^2}{\bar{S}^2} - B(s_t - \bar{s}).$$

If $B > 0$, then the log risk-free rate is countercyclical: when s is above its steady-state level, mean-reversion implies that marginal utility is expected to increase in the future, incentivizing agents to save and decreasing interest rates. Wachter (2006) shows that this condition is necessary for an upward sloping real term structure of interest rates. However, as pointed out by Verdelhan (2010), the model requires procyclical interest rates ($B < 0$) in order to generate the empirically observed relationship between interest rate differentials and currency risk premia at the short end: as equation (7) implies, there must be more priced risk in low interest rate countries than in high interest countries.

The conditional entropy of the permanent SDF component is constant:

$$L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) = \frac{\gamma^2\sigma^2}{2},$$

whereas the conditional entropies of both the SDF and the transitory SDF component are time varying, as they are functions of the log surplus consumption ratio s. The price of the long-term bond converges to $\lim_{k\to\infty} P(k)_t = \exp\left(-\gamma(s_t - \bar{s})\right)$. Finally, the permanent component of exchange rate changes is given by $\log\left(\frac{S_{t+1}^{\mathbb{P}}}{S_t^{\mathbb{P}}}\right) = -\gamma\left(\Delta c_{t+1} - \Delta c_{t+1}^*\right)$, so it is not affected by the surplus consumption ratio.

Symmetric Model with Country-Specific Shocks In a symmetric habit model (i.e., a model in which all countries share the same parameters) with country-specific shocks, long-run risk neutrality (Condition 1) is automatically satisfied: variation in the price of risk, governed by s, does not affect marginal utility and exchange rates in the long run. The long-run loading of the exchange rate on the surplus consumption ratio is given by: $\sum_{i=1}^{\infty} E_t[\Delta s_{t+i}] = \sum_{i=1}^{\infty} E_t[\log\frac{\Lambda_{t+i}}{\Lambda_t} - \log\frac{\Lambda_{t+i}^*}{\Lambda_t^*}] = -\sum_{i=1}^{\infty}\phi^{i-1}\gamma(1-\phi)(s_t^* - s_t) = -\gamma(s_t^* - s_t)$. Thus, long-run U.I.P holds,

$$\lim_{k\to\infty} E_t[\Delta s_{t\to t+k}] = -\gamma(s_t^* - s_t) = \lim_{k\to\infty} k\left(y_t^{(k),*} - y_t^{(k)}\right),$$

even though exchange rates are non-stationary in levels. A decrease in the foreign surplus consumption ratio causes foreign long-term rates to increase and the foreign currency to depreciate in the long run. The slope coefficients in regressions of Δs_{t+i} on the interest rate spread $r_t^{f,*} - r_t^f$ are equal to $-\frac{\phi^{i-1}\gamma(1-\phi)}{B}$, so they decline geometrically in absolute value as i increases, and their infinite sum equals $-\frac{\gamma}{B}$. When $B < 0$, all these slope coefficients are positive: a decrease in the foreign short rate causes the foreign currency to depreciate on average next period and all periods after that, in line with the increase in the foreign long rate. These slope coefficients cannot switch signs to match the evidence in Engel (2016), Valchev (2016), and Dahlquist and Penasse (2016).

Asymmetric model with Common Shocks We can generalize this model to N countries under the assumption of common (i.e., global) shocks and heterogeneous exposure of countries' SDFs to those shocks. In order for Condition 1 to hold, countries can only differ in their surplus consumption ratio parameters (ϕ, B), as differences in the other parameters (γ, σ^2) would imply differences in the conditional entropy of the permanent component of the SDFs, and thus differences in long-maturity bond returns expressed in the same units. Thus, Condition 1 limits the source of heterogeneity to choices leading to different real interest rate persistence across countries.

6.2 Long-Run Risks Model

We now consider the long-run risks model, proposed by Bansal and Yaron (2004) and further explored by Colacito and Croce (2011), Bansal and Shaliastovich (2013) and Engel (2016) in the context of exchange rates. In this class of models, the representative agent has utility over consumption given by:

$$\log U_t = (1 - \frac{1}{\psi}) \log \left((1-\delta) C^{1-\frac{1}{\psi}} + \delta E_t \left[U_{t+1}^{1-\gamma} \right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right),$$

where ψ represents the intertemporal elasticity of substitution in an environment without risk. Aggregate consumption growth Δc_{t+1} has a persistent component x_t, and both consumption growth shocks and shocks in

x_t exhibit conditional heteroskedasticity:

$$\begin{aligned}
\Delta c_{t+1} &= \mu + x_t + \sqrt{u_t}\varepsilon^c_{t+1}, \\
x_{t+1} &= \phi^x x_t + \sqrt{w_t}\varepsilon^x_{t+1}, \\
u_{t+1} &= (1-\phi^u)\theta^u + \phi^u u_t + \sigma^u \varepsilon^u_{t+1}, \\
w_{t+1} &= (1-\phi^w)\theta^w + \phi^w w_t + \sigma^w \varepsilon^w_{t+1}.
\end{aligned}$$

All innovations are i.i.d. standard normal. The log SDF evolves as:

$$\log \frac{\Lambda_{t+1}}{\Lambda_t} = A_0 + A_1 x_t + A_2 u_t + A_3 w_t + B_1\sqrt{u_t}\varepsilon^c_{t+1} + B_2\sqrt{w_t}\varepsilon^x_{t+1} + B_3\varepsilon^u_{t+1} + B_4\varepsilon^w_{t+1},$$

where $\{A_0, A_1, A_2, A_3, B_1, B_2, B_3, B_4\}$ are constants.[10] For convenience, we assume that the agent has preferences for early resolution of uncertainty ($\gamma > \frac{1}{\psi}$), so $B_2 < 0$. It immediately follows that conditional SDF entropy and the equilibrium log risk-free rate are given by:

$$\begin{aligned}
L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) &= \frac{1}{2} var_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{1}{2}\left({B_1}^2 u_t + {B_2}^2 w_t + {B_3}^2 + {B_4}^2\right), \\
r^f_t &= -A_0 - \frac{1}{2}\left({B_3}^2 + {B_4}^2\right) + \frac{1}{\psi}x_t - \frac{1}{2}\left(\frac{\gamma-1}{\psi} + \gamma\right)u_t - \frac{1}{2}\left(\frac{1}{\psi} - \gamma\right)\left(\frac{1}{\psi} - 1\right)\left(\frac{\kappa}{1-\kappa\phi^x}\right)^2 w_t.
\end{aligned}$$

The necessary condition (7) highlights how this model can replicate the U.I.P. puzzle: for procyclical interest rates (with respect to u_t and w_t), high interest rates correspond to low volatility SDFs.

The real bond prices in logs are affine in the state variables: $p^{i,(n)}_t = -C^{i,n}_0 - C^n_1 x_t - C^{i,n}_2 u_t - C^{i,n}_3 w_t$. In the long-run risk model, the conditional entropy of the permanent SDF component is given by:

$$L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) = \frac{1}{2} var_t\left(\log\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) = \frac{1}{2}\left({B_1}^2 u_t + (B_2 - C^\infty_1)^2 w_t + (B_3 - C^\infty_2\sigma^u)^2 + (B_4 - C^\infty_3\sigma^w)^2\right).$$

where $C^\infty_1 = \frac{1}{\psi(1-\phi^x)}$, $C^\infty_2 = -\frac{A_2 + \frac{1}{2}B_1^2}{1-\phi^u}$, and $C^\infty_3 = -\frac{A_3 + \frac{1}{2}(B_2 - C^\infty_1)^2}{1-\phi^w}$.

[10] More precisely, the constants are $A_1 = -\frac{1}{\psi}$, $A_2 = \left(\frac{1}{\psi} - \gamma\right)\frac{\gamma-1}{2}$, $A_3 = \left(\frac{1}{\psi} - \gamma\right)\frac{\gamma-1}{2}\left(\frac{\kappa}{1-\kappa\phi^x}\right)^2$, $B_1 = -\gamma$, $B_2 = \left(\frac{1}{\psi} - \gamma\right)\frac{\kappa}{1-\kappa\phi^x}$, $B_3 = \left(\frac{1}{\psi} - \gamma\right)\frac{1-\gamma}{2}\frac{\kappa}{1-\kappa\phi^u}\sigma^u$, and $B_4 = \left(\frac{1}{\psi} - \gamma\right)\frac{1-\gamma}{2}\left(\frac{\kappa}{1-\kappa\phi^x}\right)^2\frac{\kappa}{1-\kappa\phi^w}\sigma^w$, where $\kappa \equiv \frac{\delta e^{\left(1-\frac{1}{\psi}\right)\bar{m}}}{1-\delta+\delta e^{\left(1-\frac{1}{\psi}\right)\bar{m}}}$ and $\bar{m}$ is the point around which a log-linear approximation is taken (see Engel (2015) for details); if $\bar{m} = 0$, then $\kappa = \delta$.

Symmetric Model with Country-Specific Shocks The quantity of risk is governed by u_t, the volatility of consumption growth, and w_t, the volatility of expected consumption growth. Both of these forces feed into the quantity of permanent risk unless $B_1 = B_2 - C_1^{\infty} = 0$. Thus, in a symmetric LRR model (i.e., when countries share the same parameters) with country-specific shocks and heteroskedasticity, Condition 1 holds only if the model parameters satisfy the following restriction: $\gamma = 0 = \frac{1}{\psi}$, implying that the pricing kernel is constant and the investor is risk-neutral. In this case, the model counterfactually replicates the U.I.P. condition in the short-run.[11]

Asymmetric model with Common Shocks Thus, a natural extension to the model would feature common volatility processes, such that $u_t = u_t^*$ and $w_t = w_t^*$, relieving the strong parameter restriction above (see Colacito, Croce, Gavazzoni, and Ready, 2017, for a multi-country LRR model with common shocks). Condition 1 again tells us where to introduce heterogeneity in a future version of this model. For the conditional entropy of the permanent SDF component to be identical across countries, we need the following parameter restriction: $B_1 = B_1^*$ and $B_2 - C_1^{\infty} = B_2^* - C_1^{*,\infty}$. In this case, we have different SDF entropy to generate carry risk premia at the short end of the curve, but the same entropy of the permanent SDF component across countries if $B_2 - C_1^{\infty} = B_2^* - C_1^{*,\infty}$ and $C_1^{\infty} - C_1^{*,\infty} = B_2^* - B_2 \neq 0$.

These restrictions have bite. Consider an example with only heterogeneity in the persistence of the shocks. Our conditions are satisfied if $\gamma = \gamma^*$, $\delta = \delta^*$ and $\psi = \psi^*$, but $\phi^x \neq \phi^{x,*}$ such that $(1 - \delta\phi^x)(1 - \delta\phi^{x,*}) = \delta^2(1 - \gamma\psi)(1 - \phi^x)(1 - \phi^{x,*})$. That restriction cannot be satisfied when agents have a preference for early resolution of uncertainty ($\gamma\psi > 1$), as is invariably assumed in LRR models. The constant component of the entropy above adds even more parameter restrictions.

[11] In the long-run, U.I.P. is violated for risk-related innovations because the long-run loadings of the level of the exchange rate on (u_t, w_t) do not line up with the loadings of the long rates:

$$\sum_{i=1}^{\infty} E_t[\Delta s_{t+i}] = \frac{A_1}{1-\phi_x}(x_t - x_t^*) + \frac{A_2}{1-\phi_u}(u_t - u_t^*) + \frac{A_3}{1-\phi_w}(w_t - w_t^*) \neq C_1^{\infty}(x_t^* - x_t) + C_2^{\infty}(u_t^* - u_t) + C_3^{\infty}(w_t^* - w_t),$$

because $C_2^{\infty} \neq -\frac{A_2}{1-\phi_u}$ and $C_3^{\infty} \neq -\frac{A_3}{1-\phi_w}$.

6.3 Disasters Model

In the Farhi and Gabaix (2016) version of the Gabaix (2012) and Wachter (2013) rare disasters model with time-varying disaster intensity, the SDF has the following law of motion:

$$\frac{\Lambda_{t+1}}{\Lambda_t} = \frac{\Lambda^*_{t+1}}{\Lambda^*_t} \frac{\omega_{t+1}}{\omega_t} \frac{1 + Ax_{t+1}}{1 + Ax_t},$$

where Λ^* denotes the global component of marginal utility:

$$\frac{\Lambda^*_{t+1}}{\Lambda^*_t} = e^{-R} \times \begin{cases} 1, \text{if there is no disaster at } t+1 \\ B_{t+1}^{-\gamma}, \text{if there is a disaster at } t+1, \end{cases}$$

ω_{t+1} denotes the country-specific productivity:

$$\frac{\omega_{t+1}}{\omega_t} = e^{g_\omega} \times \begin{cases} 1, \text{if there is no disaster at } t+1 \\ F_{t+1}, \text{if there is a disaster at } t+1, \end{cases}$$

x_t is the (scaled by e^{-h_*}) time-varying component of the resilience of the country (with persistence ϕ_H) and $A > 0$ depends on the model parameters, one of which is the investment depreciation rate λ. After some algebra, we obtain the following expression for conditional entropy:

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \log(1 + H_t) - p_t E_t^D\left[\log(B_{t+1}^{\gamma} F_{t+1})\right] + L_t\left(\frac{1 + Ax_{t+1}}{1 + Ax_t}\right),$$

where p_t is the conditional probability of a disaster occurring next period and E_t^D is the period t expectation conditional on a disaster occurring next period. The equilibrium log risk-free rate is

$$r_t^f = (R - g_\omega - h_*) + \log\left(\frac{1 + Ax_t}{1 + (Ae^{-\phi_H} + 1)x_t}\right).$$

Thus, the risk-free rate is decreasing in x and, thus, in the resilience of the country. Again, high interest rate countries correspond to low volatility SDFs.

Finally, the conditional entropy of the permanent SDF component is

$$L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) = \log(1+H_t) - p_t E_t^D\left[\log(B_{t+1}^{\gamma} F_{t+1})\right] + L_t\left(\frac{c+x_{t+1}}{c+x_t}\right).$$

Symmetric Model with Country-Specific Shocks First, we consider a version of the model in which the parameters are the same in each country, but the shocks are country-specific. The time-varying disaster risk directly, driven partly by some country-specific shocks, affects the total quantity of permanent risk, thus violating the long-run risk neutrality Condition 1, unless the disaster intensity is constant (p_t, x_t, H_t are constant), and hence all risk premia are constant.

Asymmetric model with Common Shocks Second, we consider a version of the disaster model with common shocks, but asymmetric exposures. It is possible to introduce differences across countries that produce differences in carry trade portfolio returns at the short but not at the long end of the curve, but the heterogeneity is clearly restricted by Condition 1 to the parameters R, λ or g_w.

6.4 Factor Model

The Lustig, Roussanov, and Verdelhan (2014) model has country-specific and common shocks and arry trade risk premia arise from asymmetric exposures to global shocks. If the entropy of the permanent SDF component cannot differ across countries, then all countries' pricing kernels need the same loadings on the permanent component of the global factors. In the Lustig, Roussanov, and Verdelhan (2014) model, the permanent component of the SDF is given by:

$$\begin{aligned}\log\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}} &= \log\beta^{-1} - \alpha - \chi z_t^i - \sqrt{\gamma z_t^i}u_{t+1}^i - \tau z_t^w - \sqrt{\delta^i z_t^w}u_{t+1}^w - \sqrt{\kappa z_t^i}u_{t+1}^g \\ &\quad C_1^{\infty,\$}\left[(\phi-1)(z_t^i-\theta) - \sigma\sqrt{z_t^i}u_{t+1}^i\right] + C_2^{\infty,\$,i}\left[(\phi^w-1)(z_t^w-\theta^w) - \sigma\sqrt{z_t^w}u_{t+1}^w\right].\end{aligned}$$

The U.S. term premium is simply $E_t[rx_{t+1}^{(\infty)}] = \frac{1}{2}(\gamma z_t + \delta z_t^w)$, which is equal to one-half of the variance of the log stochastic discount factor. The foreign long bond risk premium in dollars is then simply:

$$E_t[rx_{t+1}^{(\infty),*}] + E_t[rx_{t+1}^{FX,*}] = \left[\frac{1}{2}(\gamma+\kappa)z_t + (C_1^{\infty,\$}(1-\phi)-\chi)z_t^*\right] + \left[\frac{1}{2}\delta + C_2^{\infty,\$,*}(1-\phi^w) - \tau - \eta^w\right]z_t^w,$$

where $C_1^{\infty,\$}, C_2^{\infty,\$}$ represent the loadings of the nominal long rates on the two factors. Condition 1 thus holds if $C_1^{\infty,\$}(1-\phi) = \chi$, $\kappa = 0$, and $C_2^{\infty,\$,*}(1-\phi^w) = \tau + \eta^w$. The first two restrictions rule out permanent effects of country-specific shocks, while the last restriction rules out permanent effects of global shocks (u^w). When these restrictions are satisfied, the pricing kernel is not subject to permanent shocks, and the expected foreign log holding period return on a foreign long-term bond converted into U.S. dollars is equal to the U.S. term premium: $E_t[rx_{t+1}^{(\infty),*}] + E_t[rx_{t+1}^{FX,*}] = \frac{1}{2}(\gamma z_t + \delta z_t^w)$. The higher foreign currency risk premium for investing in high δ countries is exactly offset by the lower bond risk premium. As all these models show, Proposition 1 and Condition 1 offer a simple diagnostic to assess the term structure of currency carry trade risk premia in no-arbitrage models.

The restrictions $C_1^{\infty,\$}(1-\phi) = \chi$, $\kappa = 0$, and $C_2^{\infty,\$,*,}(1-\phi^w) = \tau + \eta^w$ have a natural interpretation as restrictions on the long-run loadings of the exchange rate on the risk factors: $\sum_{i=1}^{\infty} E_t[\Delta s_{t+i}] = \sum_{i=1}^{\infty} E_t[m_{t+i} - m_{t+i}^*] = \sum_{i=1}^{\infty} \phi^{i-1}\chi(z_t^* - z_t)$. As can easily be verified, these two restrictions imply that the long-run loading of the exchange rate on the factors equals the loading of long-term interest rates:

$$\lim_{k\to\infty} E_t[\Delta s_{t\to t+k}] = \frac{\chi}{(1-\phi)}(z_t^* - z_t) = C_1^{\infty,\$}(z_t^* - z_t) = \lim_{k\to\infty} k\left(y_t^{(k),*} - y_t^{(k)}\right),$$

where we have used $C_2^{\infty,\$} = C_2^{\infty,\$,*}$. Hence, in the context of this model, our restrictions enforce long-run U.I.P.[12] In this special case, $\frac{\chi}{(1-\phi)} = C_1^{\infty} = \sqrt{\gamma}/\sigma > 0$. An increase in risk abroad causes the long rates to go up abroad and the foreign exchange rate to depreciate in the long run, but given these long-run restrictions, the initial expected exchange rate impact has to have the same sign ($\chi > 0$), thus violating the empirical evidence, as we explain below.

Our preference-free conditions constrains the sum of slope regression coefficients in a regression of future exchange rate changes Δs_{t+i} on the current interest rate spread $r_t^{f,\$,*} - r_t^{f,\$}$ to be equal to the response of long-term interest rates. Engel (2016), Valchev (2016), and Dahlquist and Penasse (2016) study these slope coefficients and find that they switch signs with the horizon i: an increase in the short-term interest rate initially cause exchange rates to appreciate, but they subsequently depreciate on average. In the factor model with a

[12] When all innovations have an impact on risk, as is the case in this model, Condition 1 rules out permanent shocks.

single country-specific factor, these slope coefficients in a regression of Δs_{t+i} on the $r_t^{f,\$,*} - r_t^{f,\$}$, given by

$$E_t \Delta s_{t+i} = \frac{\phi^{i-1}\chi}{\chi - \frac{1}{2}\gamma}\left(r_t^{f,\$,*} - r_t^{f,\$}\right),$$

decline geometrically as i increases, and their infinite sum equals $\frac{C_1^{\infty}}{\chi - \frac{1}{2}\gamma}$. When $(\chi - \frac{1}{2}\gamma) < 0$, the model can match the short-run forward premium puzzle: when the foreign short rate increases, the currency subsequently appreciates, but it continues to appreciate as long rates decline abroad. As a result, this model cannot match the sign switch in these regression coefficients.[13]

7 Additional Implications

We end this paper with two additional implications of our main results that can further help build the next generation of international finance models and guide future empirical work.

7.1 A Lower Bound on Cross-Country Correlations of the Permanent SDF Components

Brandt, Cochrane, and Santa-Clara (2006) show that the combination of relatively smooth exchange rates and much more volatile SDFs implies that SDFs are very highly correlated across countries. A 10% volatility in exchange rate changes and a volatility of marginal utility growth rates of 50% imply a correlation of at least 0.98.[14] We can derive a specific bound on the covariance of the permanent component across different countries.

[13] A richer version of the factor model with multiple country-specific risk factors can generate richer dynamics. Consider the same model with two country-specific risk factors. The long-run impulse responses of the exchange rate to short-term interest rate shocks is driven by:

$$\sum_{i=1}^{\infty} E_t[\Delta s_{t+i}] = \sum_{i=1}^{\infty} E_t[m_{t+i} - m_{t+i}^*] = \sum_{i=1}^{\infty}\left[\phi_1^{i-1}\chi_1(z_t^{1,*} - z_t^1) + \phi_2^{i-1}\chi_2(z_t^{2,*} - z_t^2)\right].$$

The slope coefficients in a regression of future exchange rate changes on the current interest rate spread $r_t^{f,\$,*} - r_t^{f,\$}$ are given by

$$E_t \Delta s_{t+i} = \frac{\phi_1^{i-1}\chi_1(\chi_1 - \frac{1}{2}\gamma_1) + \phi_2^{i-1}\chi_2(\chi_2 - \frac{1}{2}\gamma_2)}{(\chi_1 - \frac{1}{2}\gamma_1)^2 + (\chi_2 - \frac{1}{2}\gamma_2)^2}\left(r_t^{f,\$,*} - r_t^{f,\$}\right).$$

These coefficients can switch signs as we increase the maturity i if the risk factors have sufficiently heterogeneous persistence (ϕ_1, ϕ_2), and provided that $(\chi_1 - \frac{1}{2}\gamma_1)$ and $(\chi_2 - \frac{1}{2}\gamma_2)$ have opposite signs.

[14] We do not interpret the correlation of SDFs or their components in terms of cross-country risk-sharing, because doing so requires additional assumptions. The nature and magnitude of international risk sharing is an important and open question in macroeconomics (see, for example, Cole and Obstfeld (1991); Wincoop (1994); Lewis (2000); Gourinchas and Jeanne (2006); Lewis and Liu (2015); Coeurdacier, Rey, and Winant (2013); Didier, Rigobon, and Schmukler (2013); as well as Colacito and Croce (2011) and Stathopoulos (2017) on the high international correlation of state prices). A necessary but not sufficient condition to interpret the SDF correlation is for example that the domestic and foreign agents consume the same baskets of goods and participate in complete financial markets. Even in this case, the interpretation is subject to additional assumptions. In a multi-good world,

Proposition 4. *If the permanent SDF component is unconditionally lognormal, the cross-country covariance of the SDF' permanent components is bounded below by:*

$$cov\left(\log\frac{\Lambda^{\mathbb{P},*}_{t+1}}{\Lambda^{\mathbb{P},*}_{t}}, \log\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_{t}}\right) \quad\geq\quad E\left(\log\frac{R^{*}_{t+1}}{R^{(\infty),*}_{t+1}}\right) + E\left(\log\frac{R_{t+1}}{R^{(\infty)}_{t+1}}\right) - \frac{1}{2}var\left(\log\frac{S^{\mathbb{P}}_{t+1}}{S^{\mathbb{P}}_{t}}\right). \tag{19}$$

for any positive returns R_{t+1} and R^{*}_{t+1}. A conditional version of the expression holds for conditionally lognormal permanent pricing kernel components.

This result therefore extends the insights of Brandt, Cochrane, and Santa-Clara (2006) to the permanent components of the SDFs. Chabi-Yo and Colacito (2015) extend this lower bound to non-Gaussian pricing kernels and different horizons.

Since exchange rate changes and their temporary components are observable (thanks to the bonds' holding period returns), one can compute the variance of the permanent component of exchange rates, $var\left(\log\frac{S^{\mathbb{P}}_{t+1}}{S^{\mathbb{P}}_{t}}\right)$. In the data; the contribution of the last term is on the order of 1% or less. Given the large size of the equity premium compared to the term premium (a 7.5% difference according to Alvarez and Jermann, 2005), and the relatively small variance of the permanent component of exchange rates, the lower bound in Proposition 4 implies a large correlation of the permanent components.

In Figure 3, we plot the implied correlation of the permanent component against the volatility of the permanent component in the symmetric case for two different scenarios. The dotted line is for $Std\left(\log S^{\mathbb{P}}_{t}/S^{\mathbb{P}}_{t+1}\right) =$ 10%, and the plain line is for $Std\left(\log S^{\mathbb{P}}_{t}/S^{\mathbb{P}}_{t+1}\right) = 16\%$. In both cases, the implied correlation of the permanent components of the domestic and foreign pricing kernels is clearly above 0.9.

While Brandt, Cochrane, and Santa-Clara (2006) find that the SDFs are highly correlated across countries, we find that the permanent components of the SDFs, which are the main sources of volatility for the SDFs, are highly correlated across countries.

7.2 A New Long-Term Bond Return Parity Condition

Finally, we end this paper with a potential new benchmark for exchange rates. While hundreds of papers have tested the U.I.P. condition, which assumes risk-neutrality, we suggest a novel corner case, this time taking risk

variation in the relative prices of the goods drives a wedge between the pricing kernels, even in the case of perfect risk sharing (Cole and Obstfeld (1991)). Likewise, when markets are segmented, as in Alvarez, Atkeson, and Kehoe (2002) and Alvarez, Atkeson, and Kehoe (2009), the correlation of SDFs does not imply risk-sharing of the non-participating agents.

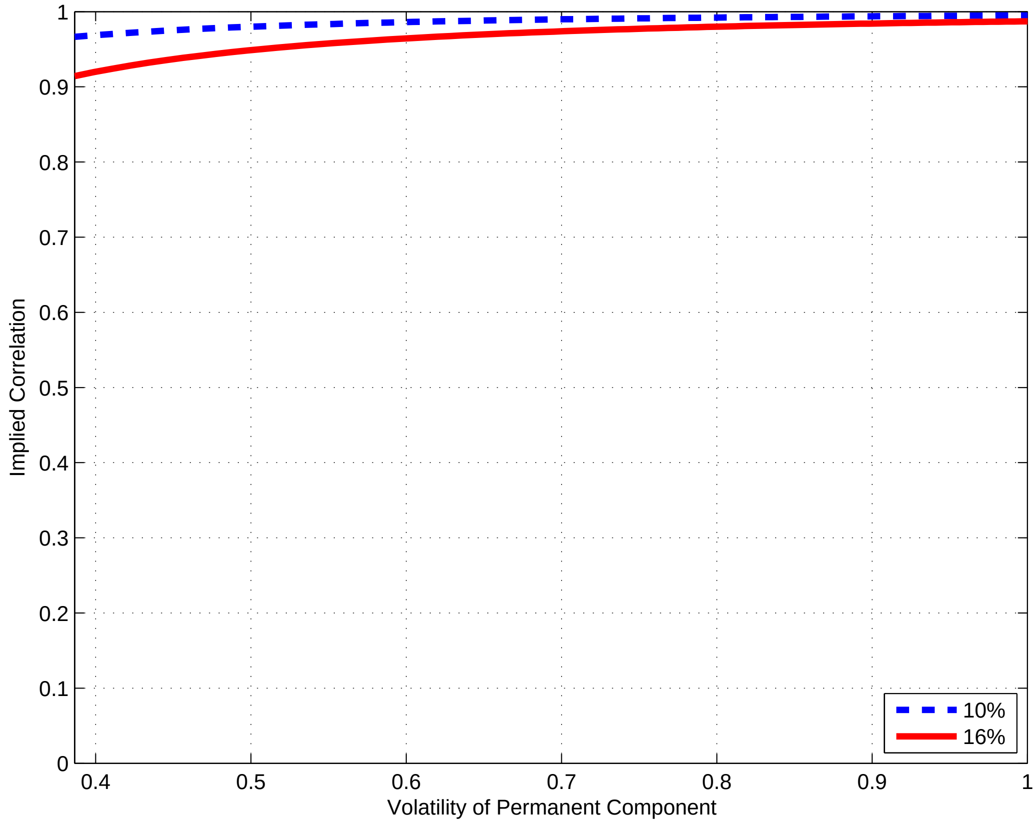

Figure 3: Cross-country Correlation of Permanent SDF Shocks — In this figure, we plot the implied correlation of the domestic and foreign permanent components of the SDF against the standard deviation of the permanent component of the SDF. The dotted line is for $Std\left(\log S_t^{\mathbb{P}}/S_{t+1}^{\mathbb{P}}\right) = 10\%$. The straight line is for $Std\left(\log S_t^{\mathbb{P}}/S_{t+1}^{\mathbb{P}}\right) = 16\%$. Following Alvarez and Jermann (2005), we assume that the equity minus bond risk premia are 7.5% in the domestic and foreign economies.

into account. When countries share permanent innovations to their SDFs, a simple long bond return parity condition emerges:

Proposition 5. *If the domestic and foreign pricing kernels have common permanent innovations, $\Lambda_{t+1}^{\mathbb{P}}/\Lambda_t^{\mathbb{P}} = \Lambda_{t+1}^{\mathbb{P},*}/\Lambda_t^{\mathbb{P},*}$ for all states, then the one-period returns on the foreign longest maturity bonds in domestic currency are identical to the domestic ones:*

$$R_{t+1}^{(\infty),*} \frac{S_t}{S_{t+1}} = R_{t+1}^{(\infty)} \text{ for all states.} \tag{20}$$

While Proposition 1 is about expected returns, Proposition 5 focuses on realized returns. In this polar case, even if most of the innovations to the pricing kernel are highly persistent, the shocks that drive exchange rates are not, because the persistent shocks are the same across countries. In that case, the exchange rate is a stationary process. In the absence of arbitrage opportunities, the currency exposure of a foreign long-term bond position to the stationary components of the pricing kernels is fully hedged by its interest rate risk exposure and does not affect the return differential with domestic bonds, which then measures the wedge between the non-stationary components of the domestic and foreign pricing kernels. When nominal exchange rates are stationary, this wedge is zero and long bond return parity obtains: bonds denominated in different currencies earn the same dollar returns, date by date.

8 Conclusion

This paper presents a simple but puzzling fact: while holding period bond returns, expressed in a common currency, differ across G10 countries at the short end of the yield curve (the U.I.P. puzzle), they are rather similar at the long end. In other words, the term structure of currency carry trade risk premia is downward-sloping. Some recent no-arbitrage models of international finance that replicates the U.I.P. puzzle fail to replicate the downward-sloping term structure of risk premia.

Under the assumption that markets are complete, we derive a simple preference-free result that helps assess the existing models and guide future theoretical and empirical work. In order to exhibit similar long-term bond returns when expressed in the same units, no arbitrage models need to exhibit the same volatility of the permanent components of their SDFs. This condition implies novel parameter restrictions in the recent no-arbitrage models of international finance.

Our results show that exchange rate risk is different from equity and bond risk. In order to account for the high equity premium and the low term premium, most of the variation in the marginal utility of wealth in no-arbitrage models must come from permanent shocks. But differences in the impact of permanent shocks to marginal utilities would lead to counterfactual differences in long-term bond risk premia across countries. Thus, in order to account for the U.I.P. puzzle and the currency carry trade risk premium at the short end of the yield curve, no-arbitrage models need to feature differences in how temporary shocks affect the marginal utility of wealth.

References

Alvarez, F., A. Atkeson, and P. Kehoe (2002): "Money, Interest Rates and Exchange Rates with Endogenously Segmented Markets," *Journal of Political Economy*, 110(1), 73–112.

——— (2009): "Time-Varying Risk, Interest Rates and Exchange Rates in General Equilibrium," *Review of Economic Studies*, 76, 851–878.

Alvarez, F., and U. J. Jermann (2004): "Using Asset Prices to Measure the Cost of Business Cycles," *Journal of Policitical Economy*, 112(6), 1123–1256.

——— (2005): "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica*, 73(6), 1977–2016.

Ang, A., and J. S. Chen (2010): "Yield Curve Predictors of Foreign Exchange Returns," Working Paper, Columbia University.

Bacchetta, P., and E. van Wincoop (2005): "Rational Inattention: A Solution to the Forward Discount Puzzle," Working Paper 11633, National Bureau of Economic Research.

Backus, D., N. Boyarchenko, and M. Chernov (2016): "Term Structures of Asset Prices and Returns," Working Paper 22162, National Bureau of Economic Research.

Backus, D., M. Chernov, and S. Zin (2014): "Sources of Entropy in Representative Agent Models," *Journal of Finance*, 69 (1), 51–99.

Backus, D., S. Foresi, and C. Telmer (2001): "Affine Models of Currency Pricing: Accounting for the Forward Premium Anomaly," *Journal of Finance*, 56, 279–304.

Backus, D. K., A. W. Gregory, and S. E. Zin (1989): "Risk Premiums in the Term Structure: Evidence from Artificial Economies," *Journal of Monetary Economics*, 24(3), 371 – 399.

Bakshi, G., and F. Chabi-Yo (2012): "Variance Bounds on the Permanent and Transitory Components of Stochastic Discount Factors," *Journal of Financial Economics*, 105, 191–208.

BANSAL, R. (1997): "An Exploration of the Forward Premium Puzzle in Currency Markets," *Review of Financial Studies*, 10, 369–403.

BANSAL, R., AND B. N. LEHMANN (1997): "Growth Optimal Portfolio Restrictions on Asset Pricing Models," *Macroeconomic Dynamics*, 1, 333–354.

BANSAL, R., AND I. SHALIASTOVICH (2013): "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 26(1), 1–33.

BANSAL, R., AND A. YARON (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

BEKAERT, G. (1996): "The Time Variation of Risk and Return in Foreign Exchange Markets: A General Equilibrium Perspective," *The Review of Financial Studies*, 9(2), 427–470.

BEKAERT, G., AND R. J. HODRICK (2001): "Expectations Hypotheses Tests," *The Journal of Finance*, 56(4), 1357–1394.

BEKAERT, G., M. WEI, AND Y. XING (2007): "Uncovered Interest Rate Parity and the Term Structure of Interest Rates," *Journal of International Money and Finance*, 26(6), 1038 – 1069.

BERGE, T. J., Ò. JORDÀ, AND A. M. TAYLOR (2011): *Currency Carry Trades*, NBER International Seminar on Macroeconomics 2010. University of Chicago Press.

BOUDOUKH, J., M. RICHARDSON, AND R. F. WHITELAW (2016): "New Evidence on the Forward Premium Puzzle," *Journal of Financial and Quantitative Analysis*, 51(3), 875–897.

BRANDT, M. W., J. H. COCHRANE, AND P. SANTA-CLARA (2006): "International Risk Sharing is Better Than You Think, or Exchange Rates are Too Smooth," *Journal of Monetary Economics*, 53(4), 671 – 698.

BRUNNERMEIER, M. K., S. NAGEL, AND L. H. PEDERSEN (2009): "Carry Trades and Currency Crashes," *NBER Macroeconomics Annual 2008*, 23, 313–347.

CAMPBELL, J. Y., AND J. H. COCHRANE (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior.," *Journal of Political Economy*, 107(2), 205–251.

Campbell, J. Y., and R. J. Shiller (1991): "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495–514.

Chabi-Yo, F., and R. Colacito (2015): "The Term Structures of Co-Entropy in International Financial Markets," Working Paper, Ohio State University.

Chernov, M., J. Graveline, and I. Zviadadze (2011): "Sources of Risk in Currency Returns," Working Paper, London School of Economics.

Chinn, M. D., and G. Meredith (2004): "Monetary Policy and Long-Horizon Uncovered Interest Parity," *IMF Staff Papers*, 51(3), 409–430.

Cochrane, J. H. (1988): "How Big Is the Random Walk in GNP?," *Journal of Political Economy*, 96(5), pp. 893–920.

Coeurdacier, N., H. Rey, and P. Winant (2013): "Financial Integration and Growth in a Risky World," Working Paper, London Business School.

Colacito, R., and M. M. Croce (2011): "Risks For The Long Run And The Real Exchange Rate," *Journal of Political Economy*, 119(1), 153–181.

Colacito, R., M. M. Croce, F. Gavazzoni, and R. C. Ready (2017): "Currency Risk Factors in a Recursive Multi-Country Economy," *Journal of Finance*, forthcoming.

Cole, H. L., and M. Obstfeld (1991): "Commodity Trade and International Risk Sharing: How Much Do Financial Markets Matter?," *Journal of Monetary Economics*, 28(1), 3–24.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53(2), 385–408.

Dahlquist, M., and H. Hasseltoft (2013): "International Bond Risk Premia," *Journal of International Economics*, 90, 12–32.

Dahlquist, M., and J. Penasse (2016): "The Missing Risk Premium in Exchange Rates," Working Paper, Stockholm School of Economics and University of Luxembourg.

Didier, T., R. Rigobon, and S. L. Schmukler (2013): "Unexploited Gains from International Diversification: Patterns of Portfolio Holdings Around the World," *Review of Economics and Statistics*, 95 (5), 1562–1583.

Dornbusch, R. (1976): "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 84(6), 1161–1176.

Driscoll, J. C., and A. C. Kraay (1998): "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *Review of Economics and Statistics*, 80(4), 549–560.

Engel, C. (1996): "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence," *Journal of Empirical Finance*, 3, 123–192.

——— (2016): "Exchange Rates, Interest Rates, and the Risk Premium," *American Economic Review*, 106(2), 436–74.

Fama, E. F. (1984): "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 14, 319–338.

Farhi, E., S. P. Fraiberger, X. Gabaix, R. Ranciere, and A. Verdelhan (2013): "Crash Risk in Currency Markets," Working Paper, Harvard University.

Farhi, E., and X. Gabaix (2016): "Rare Disasters and Exchange Rates," *The Quarterly Journal of Economics*, 131(1), 1–52.

Frankel, J. A. (1979): "On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials," *The American Economic Review*, 69(4), 610–622.

Gabaix, X. (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *The Quarterly Journal of Economics*, 127(2), 645–700.

Gabaix, X., and M. Maggiori (2015): "International Liquidity and Exchange Rate Dynamics," *The Quarterly Journal of Economics*, 130(3), 1369–1420.

Gavazzoni, F., B. Sambalaibat, and C. Telmer (2012): "Currency Risk and Pricing Kernel Volatility," Working Paper, Carnegie Mellon University.

Gourinchas, P.-O., and O. Jeanne (2006): "The Elusive Gains from International Financial Integration," *The Review of Economic Studies*, 73(3), 715–741.

Gourio, F., M. Siemer, and A. Verdelhan (2013): "International Risk Cycles," *Journal of International Economics*, 89, 471–484.

Hansen, L. P. (2012): "Dynamic Valuation Decomposition Within Stochastic Economies," *Econometrica*, 80 (3), 911–967.

Hansen, L. P., J. C. Heaton, and N. Li (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 166(2), 260–302.

Hansen, L. P., and R. J. Hodrick (1980): "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *The Journal of Political Economy*, 88(5), 829–853.

Hansen, L. P., and R. Jagannathan (1991): "Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns," *Journal of Political Economy*, 99, 225–262.

Hansen, L. P., and J. A. Scheinkman (2009): "Long-Term Risk: An Operator Approach," *Econometrica*, 77(1), 177–234.

Hassan, T. A., and R. C. Mano (2014): "Forward and Spot Exchange Rates in a Multi-Currency World," Working Paper, University of Chicago.

Joslin, S., K. J. Singleton, and H. Zhu (2011): "A New Perspective on Gaussian Dynamic Term Structure Models," *Review of Financial Studies*, 24(3), 926–970.

Jotikasthira, C., A. Le, and C. Lundblad (2015): "Why Do Term Structures in Different Currencies Co-Move?," *Journal of Financial Economics*, 115(1), 58 – 83.

Jurek, J. W. (2014): "Crash-neutral Currency Carry Trades," *Journal of Financial Economics*, 113(3), 325 – 347.

Koijen, R., T. J. Moskowitz, L. H. Pedersen, and E. B. Vrugt (2016): "Carry," *Journal of Financial Economics*, forthcoming.

Lewis, K. K. (2000): "Why Do Stocks and Consumption Imply Such Different Gains From International Risk-Sharing?," *Journal of International Economics*, 52(1), 1–35.

Lewis, K. K. (2011): "Global Asset Pricing," *Annual Review of Financial Economics*, 3, 435–466.

Lewis, K. K., and E. X. Liu (2015): "Evaluating International Consumption Risk Sharing Gains: An Asset Return View," *Journal of Monetary Economics*, 71, 84–98.

Lustig, H., N. Roussanov, and A. Verdelhan (2011): "Common Risk Factors in Currency Returns," *Review of Financial Studies*, 24(11), 3731–3777.

——— (2014): "Countercyclical Currency Risk Premia," *Journal of Financial Economics*, 111(3), 527–553.

Lustig, H., and A. Verdelhan (2007): "The Cross-Section of Foreign Currency Risk Premia and Consumption Growth Risk," *American Economic Review*, 97(1), 98–117.

——— (2015): "Does Incomplete Spanning in International Financial Markets Help to Explain Exchange Rates?," Working Paper, Stanford University and MIT.

Martin, I., and S. A. Ross (2013): "The Long Bond," Working Paper, Stanford University and MIT.

Meredith, G., and M. D. Chinn (2005): "Testing Uncovered Interest Rate Parity at Short and Long Horizons during the Post-Bretton Woods Era," Working Paper NBER 11077.

Ross, S. (2015): "The Recovery Theorem," *Journal of Finance*, 70(2), 615–648.

Stathopoulos, A. (2017): "Asset Prices and Risk Sharing in Open Economies," *The Review of Financial Studies*, 30(2), 363–415.

Valchev, R. (2016): "Bond Convenience Yields and Exchange Rate Dynamics," Working Paper, Boston College.

Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5(2), 177–188.

Verdelhan, A. (2010): "A Habit-Based Explanation of the Exchange Rate Risk Premium," *Journal of Finance*, 65 (1), 123–145.

Wachter, J. (2006): “A Consumption-Based Model of the Term Structure of Interest Rates,” *Journal of Financial Economics*, 79, 365–399.

Wachter, J. A. (2013): “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?,” *Journal of Finance*, 68(3), 987–1036.

Wincoop, E. v. (1994): “Welfare Gains From International Risk-Sharing,” *Journal of Monetary Economics*, 34(2), 175–200.

Wright, J. H. (2011): “Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset,” *American Economic Review*, 101, 1514–1534.

Wu, L. J. (2012): “Global Yield Curve Risk Factors in International Bond and Currency Portfolios,” Working Paper, UCLA.

A Theoretical Background and Proofs of Preference-Free Results

This section of the Appendix starts with a review of the Hansen and Scheinkman (2009) results and their link to the Alvarez and Jermann (2005) decomposition used in the main text. Then, we report our theoretical results on bond and currency returns in two special cases: the case of a Gaussian economy and the case of an economy with no permanent SDF shocks. The section concludes with the proofs of all the theoretical results in the main body of the paper. To make the paper self-contained, we reproduce here some proofs of intermediary results already in the literature, notably in Alvarez and Jermann (2005).

A.1 Existence and Uniqueness of Multiplicative Decomposition of the SDF

Consider a continuous-time, right continuous with left limits, strong Markov process X and the filtration $\mathcal{F}$ generated by the past values of X, completed by the null sets. In the case of infinite-state spaces, X is restricted to be a semimartingale, so it can be represented as the sum of a continuous process X^c and a pure jump process X^j. The pricing kernel process Λ is a strictly positive process, adapted to $\mathcal{F}$, for which it holds that the time t price of any payoff Π_s realized at time s ($s \geq t$) is given by

$$P_t(\Pi_s) = E\left[\frac{\Lambda_s}{\Lambda_t}\Pi_s|\mathcal{F}_t\right].$$

The pricing kernel process also satisfies $\Lambda_0 = 1$. Hansen and Scheinkman (2009) show that Λ is a multiplicative functional and establish the connection between the multiplicative property of the pricing kernel process and the semigroup property of pricing operators $\mathbb{M}$.[15] In particular, consider the family of operators $\mathbb{M}$ described by

$$\mathbb{M}_t\psi(x) = E\left[\Lambda_t\psi(X_t)|X_0 = x\right]$$

where $\psi(X_t)$ is a random payoff at t that depends solely on the Markov state at t. The family of linear pricing operators $\mathbb{M}$ satisfies $\mathbb{M}_0 = \mathbb{I}$ and $\mathbb{M}_{t+u}\psi(x) = \mathbb{M}_t\psi(x)\mathbb{M}_u\psi(x)$ and, thus, defines a semigroup, called pricing semigroup.

Further, Hansen and Scheinkman (2009) show that Λ can be decomposed as

$$\Lambda_t = e^{\beta t}\frac{\phi(X_0)}{\phi(X_t)}\Lambda_t^{\mathbb{P}}$$

where $\Lambda^{\mathbb{P}}$ is a multiplicative functional and a local martingale, ϕ is a principal (i.e. strictly positive) eigenfunction of the extended generator of $\mathbb{M}$ and β is the corresponding eigenvalue (typically negative).[16] If, furthermore, $\Lambda^{\mathbb{P}}$ is martingale, then the eigenpair (β, ϕ) also solves the principal eigenvalue problem:[17]

$$\mathbb{M}_t\phi(x) = E\left[\Lambda_t\phi(X_t)|X_0 = x\right] = e^{\beta t}\phi(x).$$

Conversely, if the expression above holds for a strictly positive ϕ and $\mathbb{M}_t\phi$ is well-defined for $t \geq 0$, then $\Lambda^{\mathbb{P}}$ is a martingale. Thus, a strictly positive solution to the eigenvalue problem above implies a decomposition

$$\Lambda_t = e^{\beta t}\frac{\phi(X_0)}{\phi(X_t)}\Lambda_t^{\mathbb{P}}$$

where $\Lambda^{\mathbb{P}}$ is guaranteed to be a martingale. The decomposition above implies that the one-period SDF is given by

$$M_{t+1} = \frac{\Lambda_{t+1}}{\Lambda_t} = e^{\beta}\frac{\phi(X_t)}{\phi(X_{t+1})}\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}$$

and satisfies

$$E\left[M_{t+1}\phi(X_{t+1})|X_t = x\right] = e^{\beta t}\phi(x).$$

Hansen and Scheinkman (2009) provide sufficient conditions for the existence of a solution to the principal eigenfunction problem and, thus, for the existence of the aforementioned pricing kernel decomposition. Notably, multiple solutions may exist,

[15] A functional Λ is multiplicative if it satisfies $\Lambda_0 = 1$ and $\Lambda_{t+u} = \Lambda_t\Lambda_u(\theta_t)$, where θ_t is a shift operator that moves the time subscript of the relevant Markov process forward by t periods. Products of multiplicative functionals are multiplicative functionals. The multiplicative property of the pricing kernel arises from the requirement for consistency of pricing across different time horizons.

[16] The extended generator of a multiplicative functional Λ is formally defined in Hansen and Scheinkman (2009) and, intuitively, assigns to a Borel function ψ a Borel function ξ such that $\Lambda_t\xi(X_t)$ is the expected time derivative of $\Lambda_t\psi(X_t)$.

[17] Since $\Lambda^{\mathbb{P}}$ is a local martingale bounded from below, it is a supermartingale. For $\Lambda^{\mathbb{P}}$ to be a martingale, additional conditions need to hold, as discussed in Appendix C of Hansen and Scheinkman (2009).

so the pricing kernel decomposition above is generally not unique. However, if the state space is finite and the Markov chain is irreducible, then Perron-Frobenious theory implies that there is a unique principal eigenvector (up to scaling), and thus a unique pricing kernel decomposition. Although multiple solutions typically exist, Hansen and Scheinkman (2009) show that the only (up to scaling) principal eigenfunction of interest for long-run pricing is the one associated with the smallest eigenvalue, as the multiplicity of solutions is eliminated by the requirement for stochastic stability of the Markov process X. In particular, only this solution ensures that the process X remains stationary and Harris recurrent under the probability measure implied by the martingale $\Lambda^{\mathbb{P}}$.

Finally, Hansen and Scheinkman (2009) show that the aforementioned pricing kernel decomposition can be useful in approximating the prices of long-maturity zero-coupon bonds. In particular, the time t price of a bond with maturity $t+k$ is given by

$$P_t^{(k)} = E\left[\frac{\Lambda_{t+k}}{\Lambda_t}|X_t = x\right] = e^{\beta k} E^{\mathbb{P}}\left[\frac{1}{\phi(X_{t+k})}|X_t = x\right]\phi(x) \approx e^{\beta k} E^{\mathbb{P}}\left[\frac{1}{\phi(X_{t+k})}\right]\phi(x)$$

where $E^{\mathbb{P}}$ is the expectation under the probability measure implied by the martingale $\Lambda^{\mathbb{P}}$ and the right-hand-side approximation becomes arbitrarily accurate as $k \to \infty$. Thus, in the limit of arbitrarily large maturity, the price of the zero-coupon bond depends on the current state solely through $\phi(x)$ and not through the expectation of the transitory component. Notably, this implies that the relevant ϕ is the one that ensures that X remains stationary under the probability measure implied by $\Lambda^{\mathbb{P}}$, i.e. the unique principal eigenfunction that implies stochastic stability for X, and β is the corresponding eigenvalue.

Indeed, Alvarez and Jermann (2005) *construct* a pricing kernel decomposition by considering a constant $\hat{\beta}$ that satisfies

$$0 < \lim_{k\to\infty} \frac{P_t^{(k)}}{\hat{\beta}^k} < \infty$$

and *defining* the transitory pricing kernel component as

$$\Lambda_t^{\mathbb{T}} = \lim_{k\to\infty} \frac{\hat{\beta}^{t+k}}{P_t^{(k)}} < \infty.$$

In contrast to Hansen and Scheinkman (2009), the decomposition of Alvarez and Jermann (2005) is constructive and not unique, as their Assumptions 1 and 2 do not preclude the existence of alternative pricing kernel decompositions to a martingale and a transitory component. Note that the Alvarez and Jermann (2005) decomposition implies that $\hat{\beta} = e^{\beta}$, where β is the smallest eigenvalue associated with a principal eigenfunction in the Hansen and Scheinkman (2009) eigenfunction problem.

A.2 Gaussian Economy

If the pricing kernel is conditionally Gaussian over horizon k, the k-horizon foreign currency risk premium satisfies:

$$E_t[rx_{t\to t+k}^{FX}] = \frac{1}{2k}\left[var_t\left(\log\frac{\Lambda_{t+k}}{\Lambda_t}\right) - var_t\left(\log\frac{\Lambda_{t+k}^*}{\Lambda_t^*}\right)\right].$$

Let us assume that the variance of the one-period SDF is constant. The annualized variance of the increase in the log SDF can be expressed as follows:

$$\frac{var\left(\log\Lambda_{t+k}/\Lambda_t\right)}{kvar\left(\Lambda_{t+1}/\Lambda_t\right)} = 1 + 2\sum_{j=1}^{k-1}\left(1-\frac{j}{k}\right)\rho_j,$$

where ρ_j denotes the j-th autocorrelation (Cochrane, 1988).[18] In the special case where the domestic and foreign countries share the same one-period volatility of the innovations, this expression for the long-run currency risk premium becomes:

$$E[rx_{t\to t+k}^{FX}] = var\left(\Delta\log\Lambda_{t+1}\right)\left[\sum_{j=1}^{k-1}\left(1-\frac{j}{k}\right)(\rho_j - \rho_j^*)\right].$$

This is the Bartlett kernel estimate with window k of the spread in the spectral density of the log SDF at zero, which measures the size of the permanent component of the SDF. More positive autocorrelation in the domestic than in the foreign pricing kernel tends to create long-term yields that are lower at home than abroad, once expressed in the same currency. The difference in yields, converted in the same units, is governed by a horse race between the speed of mean reversion in the pricing kernel at home and abroad.

[18] Cochrane (1988) uses these per period variances of the log changes in GDP to measure the size of the random walk component in GDP.

To develop some intuition for the long run, we consider the limit behavior of the foreign currency risk premium when $k \to \infty$. In the long run, the currency risk premium over many periods converges to the difference in the size of the random walk components:

$$\begin{aligned}\lim_{k\to\infty} E[rx^{FX}_{t\to t+k}] &= \frac{1}{2}var\left(\Delta \log \Lambda_{t+1}\right) \lim_{k\to\infty}\left[1+2\sum_{j=1}^{\infty}\rho_j\right] - \frac{1}{2}var\left(\Delta \log \Lambda_{t+1}\right) \lim_{k\to\infty}\left[1+2\sum_{j=1}^{\infty}\rho_j^*\right] \\ &= \frac{1}{2}\left[S_{\Delta \log \Lambda_{t+1}} - S_{\Delta \log \Lambda^*_{t+1}}\right],\end{aligned}$$

where S denotes the spectral density. The last step follows from the definition of the spectral density (see Cochrane, 1988). If the log of the exchange rate ($\log S_t$) is stationary, then the log of the foreign ($\log \Lambda^*_t$) and domestic pricing kernels ($\log \Lambda_t$) are cointegrated with co-integrating vector $(1, -1)$ and hence share the same stochastic trend component. This in turn implies that they have the same spectral density evaluated at zero. As a result, exchange rate stationarity implies that the long-run currency risk premium goes to zero.

A.3 Economy with No Permanent Innovations

Consider the special case in which the pricing kernel is not subject to permanent innovations, i.e., $\lim_{k\to\infty} \frac{E_{t+1}[\Lambda_{t+k}]}{E_t[\Lambda_{t+k}]} = 1$. For example, the Markovian environment considered by Ross (2015) to derive his recovery theorem satisfies this condition. Building on this work, Martin and Ross (2013) derive closed-form expressions for bond returns in a similar environment. Alvarez and Jermann (2005) show that this case has clear implications for domestic returns: if the pricing kernel has no permanent innovations, then the term premium on an infinite maturity bond is the largest risk premium in the economy.[19]

The absence of permanent innovations also has a strong implication for the term structure of the carry trade risk premia. When the pricing kernels do not have permanent innovations, the foreign term premium in dollars equals the domestic term premium:

$$E_t\left[rx^{(\infty),*}_{t+1}\right] + (f_t - s_t) - E_t[\Delta s_{t+1}] = E_t\left[rx^{(\infty)}_{t+1}\right].$$

The proof here is straightforward. In general, the foreign currency risk premium is equal to the difference in entropy. In the absence of permanent innovations, the term premium is equal to the entropy of the pricing kernel, so the result follows. More interestingly, a much stronger result holds in this case. Not only are the risk premia identical, but the returns on the foreign bond position are the same as those on the domestic bond position state-by-state, because the foreign bond position automatically hedges the currency risk exposure. As already noted, if the domestic and foreign pricing kernels have no permanent innovations, then the one-period returns on the longest maturity foreign bonds in domestic currency are identical to the domestic ones:

$$\lim_{k\to\infty} \frac{S_t}{S_{t+1}} \frac{R^{(k),*}_{t+1}}{R^{(k)}_{t+1}} = 1.$$

In this class of economies, the returns on long-term bonds expressed in domestic currency are equalized:

$$\lim_{k\to\infty} rx^{(k),*}_{t+1} + (f_t - s_t) - \Delta s_{t+1} = rx^{(k)}_{t+1}.$$

In countries that experience higher marginal utility growth, the domestic currency appreciates but is exactly offset by the capital loss on the bond. For example, in a representative agent economy, when the log of aggregate consumption drops more below trend at home than abroad, the domestic currency appreciates, but the real interest rate increases, because the representative agent is eager to smooth consumption. The foreign bond position automatically hedges the currency exposure.

Alvarez and Jermann (2005) propose the following example of an economy without permanent shocks: a representative agent economy with power utility investors in which the log of aggregate consumption is a trend-stationary process with normal innovations. In particular, consider the following pricing kernel (Alvarez and Jermann, 2005):

$$\log \Lambda_t = \sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i} + \beta \log t,$$

with $\epsilon \sim N(0, \sigma^2), \alpha_0 = 1$. If $\lim_{k\to\infty} \alpha_k^2 = 0$, then the SDF has no permanent component. The foreign SDF is defined similarly.

[19] If there are no permanent innovations to the pricing kernel, then the return on the bond with the longest maturity equals the inverse of the SDF: $\lim_{k\to\infty} R^{(k)}_{t+1} = \Lambda_t/\Lambda_{t+1}$. High marginal utility growth translates into higher yields on long maturity bonds and low long bond returns, and vice-versa.

In the model, the term premium equals one half of the SDF variance: $E_t\left(rx_{t+1}^{(\infty)}\right) = \sigma^2/2$, the highest possible risk premium in this economy, as the returns on the long bond are perfectly negatively correlated with the stochastic discount factor. When marginal utility is temporarily high, the representative agent would like to borrow, driving up interest rates and lowering the price of the long-term bond.

In this economy, the foreign term premium in dollars is identical to the domestic term premium:

$$E_t\left[rx_{t+1}^{(\infty),*}\right] + (f_t - s_t) - E_t[\Delta s_{t+1}] = \frac{1}{2}\sigma^2 = E_t\left[rx_{t+1}^{(\infty)}\right].$$

This result is straightforward to establish: recall that the currency risk premium is equal to the half of the difference between the domestic and the foreign SDF variance. Currencies with a high local currency term premium (high σ^2) also have an offsetting negative currency risk premium, while those with a small term premium have a large currency risk premium. Hence, U.S. investors receive the same dollar premium on foreign as on domestic bonds. There is no point in chasing high term premia around the world, at least not in economies with only temporary innovations to the pricing kernel. Currencies with the highest local term premia also have the lowest (i.e., most negative) currency risk premia.

A.4 Proofs

- Proof of Proposition 1:

  Proof. The proof builds on some results in Backus, Foresi, and Telmer (2001) and Alvarez and Jermann (2005). Specifically, Backus, Foresi, and Telmer (2001) show that the foreign currency risk premium is equal to the difference between domestic and foreign total SDF entropy:

$$(f_t - s_t) - E_t[\Delta s_{t+1}] = L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) - L_t\left(\frac{\Lambda^*_{t+1}}{\Lambda^*_t}\right).$$

  Furthermore, Alvarez and Jermann (2005) establish that total SDF entropy equals the sum of the entropy of the permanent SDF component and the expected log term premium:

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) + E_t\left(\log\frac{R_{t+1}^{(\infty)}}{R_t^f}\right).$$

  Applying the Alvarez and Jermann (2005) decomposition to the Backus, Foresi, and Telmer (2001) expression yields the desired result.

  To derive the Backus, Foresi, and Telmer (2001) expression, consider a foreign investor who enters a forward position in the currency market with payoff $S_{t+1} - F_t$. The investor's Euler equation is:

$$E_t\left(\frac{\Lambda^*_{t+1}}{\Lambda^*_t}(S_{t+1} - F_t)\right) = 0.$$

  In the presence of complete, arbitrage-free international financial markets, exchange rate changes equal the ratio of the domestic and foreign stochastic discount factors:

$$\frac{S_{t+1}}{S_t} = \frac{\Lambda_{t+1}}{\Lambda_t}\frac{\Lambda^*_t}{\Lambda^*_{t+1}},$$

  Dividing the investor's Euler equation by S_t and applying the no arbitrage condition, the forward discount is:

$$f_t - s_t = \log E_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) - \log E_t\left(\frac{\Lambda^*_{t+1}}{\Lambda^*_t}\right).$$

  The second component of the currency risk premium is expected foreign appreciation; applying logs and conditional expectations to the no arbitrage condition above leads to:

$$E_t[\Delta s_{t+1}] = E_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) - E_t\left(\log\frac{\Lambda^*_{t+1}}{\Lambda^*_t}\right).$$

Combining the two terms of the currency risk premium leads to:

$$(f_t - s_t) - E_t[\Delta s_{t+1}] = \log E_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) - E_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) - \log E_t\left(\frac{\Lambda^*_{t+1}}{\Lambda^*_t}\right) + E_t\left(\log\frac{\Lambda^*_{t+1}}{\Lambda^*_t}\right)$$

Applying the definition of conditional entropy in the equation above yields the Backus, Foresi, and Telmer (2001) expression.

To derive the Alvarez and Jermann (2005) result, first note that since the permanent component of the pricing kernel is a martingale, its conditional entropy can be expressed as follows:

$$L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) = -E_t\left(\log\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right).$$

The definition of conditional entropy implies the following decomposition of total SDF entropy:

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \log E_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) - E_t\left(\log\frac{\Lambda^{\mathbb{T}}_{t+1}}{\Lambda^{\mathbb{T}}_t}\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right)$$

or, using the above expression for the conditional entropy of the permanent SDF component:

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = -\log R^f_t - E_t\left(\log\frac{\Lambda^{\mathbb{T}}_{t+1}}{\Lambda^{\mathbb{T}}_t}\right) + L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right).$$

The Alvarez and Jermann (2005) result hinges on:

$$\lim_{k\to\infty} R^{(k)}_{t+1} = \Lambda^{\mathbb{T}}_t/\Lambda^{\mathbb{T}}_{t+1}.$$

Under the assumption that $0 < \lim_{k\to\infty}\frac{P^{(k)}_t}{\delta^k} < \infty$ for all t, one can write:

$$\lim_{k\to\infty} R^{(k)}_{t+1} = \lim_{k\to\infty}\frac{E_{t+1}\left(\frac{\Lambda_{t+k}}{\Lambda_{t+1}}\right)}{E_t\left(\frac{\Lambda_{t+k}}{\Lambda_t}\right)} = \frac{\lim_{k\to\infty}\frac{E_{t+1}(\Lambda_{t+k}/\delta^{t+k})}{\Lambda_{t+1}}}{\lim_{k\to\infty}\frac{E_t(\Lambda_{t+k}/\delta^{t+k})}{\Lambda_t}} = \frac{\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda_{t+1}}}{\frac{\Lambda^{\mathbb{P}}_t}{\Lambda_t}} = \Lambda^{\mathbb{T}}_t/\Lambda^{\mathbb{T}}_{t+1}.$$

Thus, the infinite-maturity bond is exposed only to transitory SDF risk. □

- Proof of Proposition 2:

Proof. Note that:

$$E[s^{\mathbb{T}}_t] = s_0 + t(\log\delta - \log\delta^*) + \lim_{k\to\infty} kE\left(y^{(k)}_t - y^{(k),*}_t\right) - \lim_{k\to\infty} k\left(y^{(k)}_0 - y^{(k),*}_0\right).$$

Hence, due to the stationarity of $s^{\mathbb{T}}$, we have

$$\lim_{k\to\infty} E_t[s^{\mathbb{T}}_{t+k}] = \lim_{k\to\infty} E[s^{\mathbb{T}}_{t+k}] = s_0 + \lim_{k\to\infty}(t+k)(\log\delta - \log\delta^*) + \lim_{k\to\infty} kE\left(y^{(k)}_{t+k} - y^{(k),*}_{t+k}\right) - \lim_{k\to\infty} k\left(y^{(k)}_0 - y^{(k),*}_0\right).$$

This, in turn, implies that:

$$\lim_{k\to\infty} E_t[\Delta s^{\mathbb{T}}_{t\to t+k}] = \lim_{k\to\infty} E_t[s^{\mathbb{T}}_{t+k}] - s^{\mathbb{T}}_t = \lim_{k\to\infty} k(\log\delta - \log\delta^*) + \lim_{k\to\infty} kE\left(y^{(k)}_{t+k} - y^{(k),*}_{t+k}\right) - \lim_{k\to\infty} k\left(y^{(k)}_t - y^{(k),*}_t\right),$$

or

$$\lim_{k\to\infty} E_t[\Delta s^{\mathbb{T}}_{t\to t+k}] = \lim_{k\to\infty} E_t[s^{\mathbb{T}}_{t+k}] - s^{\mathbb{T}}_t = -\lim_{k\to\infty} k\left(y^{(k)}_t - y^{(k),*}_t\right),$$

where we have used $\log\delta = -\lim_{k\to\infty} E\left(y^{(k)}_{t+k}\right)$. □

- Proof of Proposition 3:

Proof. Note that $L(x_{t+1}) = EL_t(x_{t+1}) + L_t(E(x_{t+1}))$. Given the stationary of the stochastic discount factor, $\lim_{k\to\infty}(1/k)L_t\left(E\frac{\Lambda_{t+k}}{\Lambda_t}\right) = 0$. Hence $\lim_{k\to\infty}(1/k)L\left(E\frac{\Lambda_{t+k}}{\Lambda_t}\right) = \lim_{k\to\infty}(1/k)EL_t\left(E\frac{\Lambda_{t+k}}{\Lambda_t}\right)$. Given our assumptions, it can be shown directly that:

$$\lim_{k\to\infty}(1/k)\left[L\left(\frac{\Lambda_{t+k}}{\Lambda_t}\right) - L\left(\frac{\Lambda^*_{t+k}}{\Lambda^*_t}\right)\right] = \left[L\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) - L\left(\frac{\Lambda^{\mathbb{P},*}_{t+1}}{\Lambda^{\mathbb{P},*}_t}\right)\right].$$

This result follows directly from the Alvarez-Jermann decomposition of the pricing kernel (see Alvarez and Jermann (2005)'s proposition 6).
We offer a different, shorter proof that directly exploits the Hansen-Scheinkman decomposition of the pricing kernel, but does not rely on the Alvarez-Jermann representation:

$$\frac{\Lambda_{t+k}}{\Lambda_t} = \frac{\Lambda^{\mathbb{P}}_{t+k}}{\Lambda^{\mathbb{P}}_t}\frac{\phi(X_t)}{\phi(X_{t+k})e^{\beta k}}.$$

Using a change of measure by exploiting the martingale property of the permanent component,

$$\lim_{k\to\infty}(1/k)\log E\frac{\Lambda^{\mathbb{P}}_{t+k}}{\Lambda^{\mathbb{P}}_t}\frac{\phi(X_t)}{\phi(X_{t+k})e^{\beta k}} = \lim_{k\to\infty}(1/k)\log\widehat{E}\frac{\phi(X_t)}{\phi(X_{t+k})e^{\beta k}} = 1,$$

where the last step follows because X is a Markov process. We also know that

$$\lim_{k\to\infty}(1/k)E\log\frac{\Lambda^{\mathbb{P}}_{t+k}}{\Lambda^{\mathbb{P}}_t}\frac{\phi(X_t)}{\phi(X_{t+k})e^{\beta k}} = E\log\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}.$$

As a result, we know that

$$\lim_{k\to\infty}(1/k)L\left(\frac{\Lambda_{t+k}}{\Lambda_t}\right) = 1 - E\log\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t} = L\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right).$$

□

The same result can be derived in the framework of Alvarez and Jermann (2005), building on their Proposition 5. Alvarez and Jermann (2005) show that, when the limits of the k-period bond risk premium and the yield difference between the k-period discount bond and the one-period riskless bond (when the maturity k tends to infinity) are well defined and the unconditional expectations of holding returns are independent of calendar time, then the average term premium $E\left[\lim_{k\to\infty} rx^{(k),*}_{t+1}\right]$ equals the average yield spread $E[\lim_{k\to\infty} y^{(k),*}_t - y^{(1),*}_t]$. Substituting for the term premiums in Proposition 1 leads to:

$$E[y^{(\infty),*}_t] + E[\Delta s_{t+1}] = E[y^{(\infty),*}_t] + E\left[L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) - L_t\left(\frac{\Lambda^{\mathbb{P},*}_{t+1}}{\Lambda^{\mathbb{P},*}_t}\right)\right].$$

Under regularity conditions, in a stationary environment, $E[\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}\Delta s_{t+j}] = \lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}E[\Delta s_{t+j}]$ converges to $E[\Delta s_{t+1}]$. Using this result also produces the proposition.

- Proof of Proposition 4:
  First, following Alvarez and Jermann (2005), we establish a bound for the unconditional entropy of the permanent SDF component:

$$L\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) \geq E\left(\log R_{t+1}\right) - E\left(\log R^{(\infty)}_{t+1}\right).$$

Applying unconditional expectations on the two sides of the previously established conditional SDF entropy bound leads to:

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) \geq E_t\left(\log R_{t+1}\right) - \log R^f_t,$$

using the following property of entropy: for any admissible random variable X, it holds that

$$E\left[L_t(X_{t+1})\right] = L(X_{t+1}) - L\left[E_t(X_{t+1})\right].$$

After some algebra, the following bound for the unconditional entropy of the SDF is obtained:

$$L\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) \geq L\left(\frac{1}{R_t^f}\right) + E\left(\log\frac{R_{t+1}}{R_t^f}\right).$$

To derive an expression for the unconditional entropy of the permanent SDF component, we need to decompose the unconditional SDF entropy. To do so, we start with the decomposition of the conditional SDF entropy,

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) + E_t\left(\log R_{t+1}^{(\infty)}\right) - \log R_t^f,$$

and apply unconditional expectations on both sides of the expression in order to obtain

$$L\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = L\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) + L\left(\frac{1}{R_t^f}\right) + E\left(\log\frac{R_{t+1}^{(\infty)}}{R_t^f}\right),$$

using the fact that the permanent component of the pricing kernel is a martingale:

$$L\left(E_t\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) = 0.$$

Using the above decomposition of unconditional SDF entropy, the unconditional entropy bound can be written as follows:

$$L\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) + L\left(\frac{1}{R_t^f}\right) + E\left(\log\frac{R_{t+1}^{(\infty)}}{R_t^f}\right) \geq L\left(\frac{1}{R_t^f}\right) + E\left(\log\frac{R_{t+1}}{R_t^f}\right).$$

The Alvarez and Jermann (2005) unconditional bound follows immediately by rearranging the terms in the expression above. Considering the unconditional covariance of the domestic and foreign permanent SDF components and using this bound yields the unconditional expression of Proposition 4:

$$cov\left(\log\frac{\Lambda_{t+1}^{\mathbb{P},*}}{\Lambda_t^{\mathbb{P},*}}, \log\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) \geq E\left(\log\frac{R_{t+1}^*}{R_{t+1}^{(\infty),*}}\right) + E_t\left(\log\frac{R_{t+1}}{R_{t+1}^{(\infty)}}\right) - \frac{1}{2}var\left(\log\frac{S_{t+1}^{\mathbb{P}}}{S_t^{\mathbb{P}}}\right).$$

This result relies on the assumption that $\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}$ and $\frac{\Lambda_{t+1}^{\mathbb{P},*}}{\Lambda_t^{\mathbb{P},*}}$ are unconditionally lognormal and the entropy property

$$L(X) = \frac{1}{2}var(\log X)$$

for any lognormal random variable X.

- Proof of Proposition 5:

Proof. As shown inAlvarez and Jermann (2005) (see the proof of Proposition 1), the return of the infinite maturity bond reflects the transitory SDF component:

$$\lim_{k\to\infty} R_{t+1}^{(k)} = \Lambda_t^{\mathbb{T}}/\Lambda_{t+1}^{\mathbb{T}}.$$

The result of Proposition 5 follows directly from the no-arbitrage expression for the spot exchange rate when markets are complete:

$$\frac{S_{t+1}}{S_t} = \frac{\Lambda_{t+1}}{\Lambda_t}\frac{\Lambda_t^*}{\Lambda_{t+1}^*}.$$

In this case,

$$\lim_{k\to\infty}\frac{S_t}{S_{t+1}}\frac{R_{t+1}^{(k),*}}{R_{t+1}^{(k)}} = \frac{S_t}{S_{t+1}}\frac{\lim_{k\to\infty} R_{t+1}^{(k),*}}{\lim_{k\to\infty} R_{t+1}^{(k)}} = \frac{S_t}{S_{t+1}}\frac{\Lambda_t^{\mathbb{T}}}{\Lambda_{t+1}^{\mathbb{T}}}\frac{\Lambda_{t+1}^{\mathbb{T},*}}{\Lambda_t^{\mathbb{T},*}} = \frac{\Lambda_{t+1}^{\mathbb{P},*}}{\Lambda_t^{\mathbb{P},*}}\frac{\Lambda_t^{\mathbb{P}}}{\Lambda_{t+1}^{\mathbb{P}}} = \frac{S_t^{\mathbb{P}}}{S_{t+1}^{\mathbb{P}}},$$

using the decomposition of exchange rate changes into a permanent and a transitory component:

$$\frac{S_{t+1}}{S_t} = \left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\frac{\Lambda_t^{\mathbb{P},*}}{\Lambda_{t+1}^{\mathbb{P},*}}\right)\left(\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}\frac{\Lambda_t^{\mathbb{T},*}}{\Lambda_{t+1}^{\mathbb{T},*}}\right) = \frac{S_{t+1}^{\mathbb{P}}}{S_t^{\mathbb{P}}}\frac{S_{t+1}^{\mathbb{T}}}{S_t^{\mathbb{T}}}.$$

The exposure of the domestic and foreign infinite-maturity bonds to transitory risk fully offsets the transitory component of exchange rate changes, so only the exposure to the permanent part remains. □

B Lustig, Roussanov, and Verdelhan (2014) Factor Model

This section provides details on the properties of bond and currency premia in the Lustig, Roussanov, and Verdelhan (2014) model.

SDF Decomposition The log nominal bond prices are affine in the state variable z and z^w: $p_t^{i,(n)} = -C_0^{i,n} - C_1^n z_t - C_2^{i,n} z_t^w$. To calculate the parameter set $(C_0^{i,n}, C_1^{i,n}, C_2^{i,n})$, we follow the usual recursive process. In particular, the price of a one-period nominal bond is:

$$P^{i,(1)} = E_t(M_{t+1}^i) = E_t\left(e^{-\alpha - \chi z_t - \tau z_t^w - \sqrt{\gamma z_t^i} u_{t+1}^i - \sqrt{\delta^i z_t^w} u_{t+1}^w - \sqrt{\kappa z_t^i} u_{t+1}^g - \pi_0 - \eta^w z_t^w - \sigma_\pi \epsilon_{t+1}^i}\right).$$

As a result, $C_0^1 = \alpha + \pi_0 - \frac{1}{2}\sigma_\pi^2$, $C_1^1 = \chi - \frac{1}{2}(\gamma + \kappa)$, and $C_2^{i,1} = \tau - \frac{1}{2}\delta^i + \eta^w$.

The rest of the bond prices are calculated recursively using the Euler equation: $P_t^{i,(n)} = E_t(M_{t+1}^{i,\$} P_{t+1}^{i,(n-1)})$. This leads to the following difference equations:

$$\begin{aligned}
-C_0^{i,n} - C_1^n z_t - C_2^{i,n} z_t^w &= -\alpha - \chi z_t - \tau z_t^w - C_0^{n-1} - C_1^{n-1}\left[(1-\phi)\theta + \phi z_t\right] - C_2^{i,n-1}\left[(1-\phi^w)\theta^w + \phi^w z_t^w\right] \\
&+ \frac{1}{2}(\gamma+\kappa) z_t + \frac{1}{2}\left(C_1^{n-1}\right)^2 \sigma^2 z_t - \sigma\sqrt{\gamma} C_1^{n-1} z_t \\
&+ \frac{1}{2}\delta^i z_t^w + \frac{1}{2}\left(C_2^{i,n-1}\right)^2 (\sigma^w)^2 z_t^w - \sigma^w \sqrt{\delta^i} C_2^{i,n-1} z_t^w \\
&- \pi_0 - \eta^w z_t^w + \frac{1}{2}\sigma_\pi^2
\end{aligned}$$

Solving the equations above, we recover the set of bond price parameters:

$$\begin{aligned}
C_0^{i,n} &= \alpha + \pi_0 - \frac{1}{2}\sigma_\pi^2 + C_0^{n-1} + C_1^{n-1}(1-\phi)\theta + C_2^{i,n-1}(1-\phi^w)\theta^w, \\
C_1^n &= \chi - \frac{1}{2}(\gamma+\kappa) + C_1^{n-1}\phi - \frac{1}{2}\left(C_1^{n-1}\right)^2\sigma^2 + \sigma\sqrt{\gamma} C_1^{n-1} \\
C_2^{i,n} &= \tau - \frac{1}{2}\delta^i + \eta^w + C_2^{i,n-1}\phi^w - \frac{1}{2}\left(C_2^{i,n-1}\right)^2 (\sigma^w)^2 + \sigma^w\sqrt{\delta^i} C_2^{i,n-1}.
\end{aligned}$$

The temporary pricing component of the pricing kernel is:

$$\Lambda_t^{\mathbb{T}} = \lim_{n\to\infty} \frac{\beta^{t+n}}{P_t^n} = \lim_{n\to\infty} \beta^{t+n} e^{C_0^{i,n} + C_1^n z_t + C_2^{i,n} z_t^w},$$

where the constant β is chosen in order to satisfy Assumption 1 in Alvarez and Jermann (2005): $0 < \lim_{n\to\infty} \frac{P_t^n}{\beta^n} < \infty$. The temporary pricing component of the SDF is thus equal to:

$$\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}} = \beta e^{C_1^\infty (z_{t+1} - z_t) + C_2^{i,\infty}(z_{t+1}^w - z_t^w)} = \beta e^{C_1^\infty\left[(\phi-1)(z_t^i - \theta) - \sigma\sqrt{z_t^i} u_{t+1}^i\right] + C_2^{i,\infty}\left[(\phi^w - 1)(z_t^w - \theta^w) - \sigma\sqrt{z_t^w} u_{t+1}^w\right]}.$$

The martingale component of the SDF is then:

$$\begin{aligned}
\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}} &= \frac{\Lambda_{t+1}}{\Lambda_t}\left(\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}\right)^{-1} = \beta^{-1} e^{-\alpha - \chi z_t^i - \sqrt{\gamma z_t^i} u_{t+1}^i - \tau z_t^w - \sqrt{\delta^i z_t^w} u_{t+1}^w - \sqrt{\kappa z_t^i} u_{t+1}^g} \\
&\quad e^{C_1^\infty\left[(\phi-1)(z_t^i - \theta) - \sigma\sqrt{z_t^i} u_{t+1}^i\right] + C_2^{i,\infty}\left[(\phi^w - 1)(z_t^w - \theta^w) - \sigma\sqrt{z_t^w} u_{t+1}^w\right]}.
\end{aligned}$$

As a result, we need $\chi = C_1^\infty(1-\phi)$ to make sure that the country-specific factor does not contribute a martingale component. This special case corresponds to the absence of permanent shocks to the SDF: when $C_1^\infty(1-\phi) = \chi$ and $\kappa = 0$, the permanent component of the stochastic discount factor is constant. To see this result, let us go back to the implicit definition of B_1^∞ in Equation (26):

$$
\begin{aligned}
0 &= -\frac{1}{2}(\gamma+\kappa) - \frac{1}{2}\left(C_1^\infty\right)^2\sigma^2 + \sigma\sqrt{\gamma}C_1^\infty \\
0 &= \left(\sigma C_1^\infty - \sqrt{\gamma}\right)^2,
\end{aligned}
$$

where we have imposed $\kappa = 0$. In this special case, $C_1^\infty = \sqrt{\gamma}/\sigma$. Using this result in Equation (26), the permanent component of the SDF reduces to:

$$
\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}} = \frac{\Lambda_{t+1}}{\Lambda_t}\left(\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}\right)^{-1} = \beta^{-1}e^{-\tau z_t^w - \sqrt{\delta^i z_t^w}u_{t+1}^w}e^{C_2^{i,\infty}\left[(\phi^w-1)(z_t^w-\theta^w)-\sigma\sqrt{z_t^w}u_{t+1}^w\right]}.
$$

Bond Premia The expected log excess return on a zero coupon bond is thus given by:

$$
E_t[rx_{t+1}^{(n)}] = [-\frac{1}{2}\left(C_1^{n-1}\right)^2\sigma^2 + \sigma\sqrt{\gamma}C_1^{n-1}]z_t + [-\frac{1}{2}\left(C_2^{i,n-1}\right)^2\sigma^2 + \sigma\sqrt{\delta}^i C_2^{i,n-1}]z_t^w.
$$

The expected log excess return of an infinite maturity bond is then:

$$
E_t[rx_{t+1}^{(\infty)}] = [-\frac{1}{2}\left(C_1^\infty\right)^2\sigma^2 + \sigma\sqrt{\gamma}C_1^\infty]z_t + [-\frac{1}{2}\left(C_2^{i,\infty}\right)^2\sigma^2 + \sigma\sqrt{\delta}^i C_2^{i,\infty}]z_t^w.
$$

The $-\frac{1}{2}\left(C_1^\infty\right)^2\sigma^2$ is a Jensen term. The term premium is driven by $\sigma\sqrt{\gamma}C_1^\infty z_t$, where C_1^∞ is defined implicitly in the second order equation $B_1^\infty = \chi - \frac{1}{2}(\gamma+\kappa) + C_1^\infty\phi - \frac{1}{2}\left(C_1^\infty\right)^2\sigma^2 + \sigma\sqrt{\gamma}C_1^\infty$. Consider the special case of $C_1^\infty(1-\phi) = \chi$ and $\kappa = 0$ and $C_2^{i,\infty}(1-\phi) = \tau$. In this case, the expected term premium is simply $E_t[rx_{t+1}^{(\infty)}] = \frac{1}{2}(\gamma z_t + \delta z_t^w)$, which is equal to one-half of the variance of the log stochastic discount factor.

Currency Premia The expected log excess return of the infinite maturity bond of country i is:

$$
E_t[rx_{t+1}^{(\infty),i}] = \left[C_1^\infty(1-\phi) - \chi + \frac{1}{2}(\gamma+\kappa)\right]z_t^i + \left[C_2^{i,\infty}(1-\phi^w) - \tau + \frac{1}{2}\delta^i - \eta^w\right]z_t^w.
$$

The foreign currency risk premium is given by:

$$
E_t[rx_{t+1}^{FX,i}] = -\frac{1}{2}(\gamma+\kappa)(z_t^i - z_t) + \frac{1}{2}(\delta-\delta^i)(z_t^w).
$$

Investors obtain high foreign currency risk premia when investing in currencies whose exposure to the global shocks is smaller. That is the source of short-term carry trade risk premia. The foreign bond risk premium in dollars is simply given by the sum of the two expressions above:

$$
\begin{aligned}
E_t[rx_{t+1}^{(\infty),i}] + E_t[rx_{t+1}^{FX,i}] &= \left[\frac{1}{2}(\gamma+\kappa)z_t + (C_1^\infty(1-\phi)-\chi)z_t^i\right] \\
&+ \left[\frac{1}{2}\delta + C_2^{i,\infty}(1-\phi^w) - \tau - \eta^w\right]z_t^w.
\end{aligned}
$$

Simulation Results We simulate the Lustig, Roussanov, and Verdelhan (2014) model, obtaining a panel of $T = 33600$ monthly observations and $N = 30$ countries. The calibration parameters are reported in Table 3 and the simulation results in Table 4. Each month, the 30 countries are ranked by their interest rates (Section I) or by the slope of the yield curves (Section II) into six portfolios. Low interest rate currencies on average have higher exposure δ to the world factor. As a result, these currencies appreciate in case of an adverse world shocks. Long positions in these currencies earn negative excess returns (rx^{fx}) of -4.09% per annum. On the other hand, high interest rate currencies typically have high δ. Long positions in these currencies earn positive excess returns (rx^{FX}) of 2.35% per annum. At the short end, the carry trade strategy, which goes long in the sixth portfolio and short in the first one, delivers an excess return of 6.45% and a Sharpe ratio of 0.54.

This spread is not offset by higher local currency bond risk premia in the low interest rate countries with higher δ. The log excess return on the 30-year zero coupon bond is 0.67% in the first portfolio compared to 0.97 % in the last portfolio. At the 30-year maturity, the high-minus-low carry trade strategy still delivers a profitable excess return of 6.75% and a Sharpe ratio of 0.50. This large currency risk premium at the long end of the curve stands in stark contrast to the data. Similar results obtain when sorting countries by the slopes of their yield curves. Countries with flat yield curves tend to be countries with high short-term interest rates, while countries with steep yield curves tend to be countries with low short-term interest rates. As a result, the currency carry trade is long the last portfolio in Section II and short the first portfolio. At the 30-year maturity, the carry trade strategy still delivers a profitable excess return of 6.18% and a Sharpe ratio of 0.46.

Our theoretical results help explain the shortcomings of this simulation. In Lustig, Roussanov, and Verdelhan (2014) calibration, the conditions for long run bond parity are not satisfied. First, global shocks have permanent effects in all countries, because $C_2^{i,\infty}(1-\phi^w) < \tau + \eta^w$ for all $i = 1, \ldots, 30$. Second, the global shocks are not symmetric, because δ varies across countries. The heterogeneity in δ's across countries generates substantial dispersion in exposure to the permanent component. As a result, our long-run uncovered bond parity condition is violated.

Table 3: Parameter Estimates

Stochastic discount factor					
α (%)	χ	τ	γ	κ	δ
0.76	0.89	0.06	0.04	2.78	0.36
State variable dynamics					
ϕ	θ (%)	σ (%)	ϕ^w	θ^w (%)	σ^w (%)
0.91	0.77	0.68	0.99	2.09	0.28
Inflation dynamics				Heterogeneity	
η^w	π_0 (%)	σ^π (%)		δ_h	δ_l
0.25	−0.31	0.37		0.22	0.49
Implied SDF dynamics					
$E(Std_t(\widetilde{m}))$	$Std(Std_t(\widetilde{m}))$ (%)	$E(Corr(\widetilde{m}_{t+1}, \widetilde{m}^i_{t+1}))$		$Std(z)$ (%)	$Std(z^w)$ (%)
0.59	4.21	0.98		0.50	1.32

Notes: This table reports the parameter values for the estimated version of the model. The model is defined by the following set of equations:

$$
\begin{aligned}
-\widetilde{m}^i_{t+1} &= \alpha + \chi z^i_t + \sqrt{\gamma z^i_t} u^i_{t+1} + \tau z^w_t + \sqrt{\delta^i z^w_t} u^w_{t+1} + \sqrt{\kappa z^i_t} u^g_{t+1}, \\
z^i_{t+1} &= (1-\phi)\theta + \phi z^i_t - \sigma \sqrt{z^i_t} u^i_{t+1}, \\
z^w_{t+1} &= (1-\phi^w)\theta^w + \phi^w z^w_t - \sigma^w \sqrt{z^w_t} u^w_{t+1}, \\
\pi^i_{t+1} &= \pi_0 + \eta^w z^w_t + \sigma_\pi \epsilon^i_{t+1}.
\end{aligned}
$$

All countries share the same parameter values except for δ^i, which is distributed uniformly on $[\delta_h, \delta_l]$. The home country exhibits the average δ, which is equal to 0.36.

Table 4: Simulated Excess Returns on Carry Strategies in the Lustig, Roussanov, and Verdelhan (2014) Model

	Low	2	3	4	5	High
Section I: Sorting by Interest Rate Levels						
Panel A: Exchange Rates, Interest Rates, and Bond Returns						
Δs	1.93	0.79	0.44	0.06	-0.16	-0.85
$\sigma_{\Delta s}$	11.04	9.55	9.06	8.98	9.02	9.54
$r^{f,*} - r^f$	-2.16	-1.21	-0.63	-0.10	0.43	1.50
$rx^{(30),*}$	0.67	0.75	0.79	0.89	0.93	0.97
Panel B: Carry Returns with Short-Term Bills						
rx^{FX}	-4.09	-2.00	-1.06	-0.16	0.59	2.35
Panel C: Carry Returns with Long-Term Bonds						
$rx^{(30),\$}$	-3.42	-1.25	-0.27	0.72	1.52	3.33
Section II: Sorting by Interest Rate Slopes						
Panel A: Exchange Rates, Interest Rate Slopes, and Bond Returns						
Δs	-2.06	-1.12	-0.49	-0.03	0.50	1.92
$\sigma_{\Delta s}$	11.35	9.60	8.97	8.84	8.95	9.93
$y^{10} - y^{1/4}$	-0.87	-0.42	-0.13	0.12	0.38	1.03
$rx^{(30),*}$	0.87	0.87	0.86	0.87	0.86	0.84
Panel B: Carry Returns with Short-Term Bills						
rx^{FX}	3.23	1.78	0.83	0.08	-0.76	-2.92
Panel C: Carry Returns with Long-Term Bonds						
$rx^{(30),\$}$	4.09	2.65	1.69	0.94	0.11	-2.09

Notes: The table reports summary statistics on simulated data from the Lustig, Roussanov, and Verdelhan (2014) model. Data are obtained from a simulated panel with $T = 33600$ monthly observations and $N = 30$ countries. In Section I, countries are sorted by interest rates into six portfolios. In Section II, they are sorted by the slope of their yield curves (defined as the difference between the 10-year yield and the three-month yield). In each section, Panel A reports the average change in exchange rate (Δs), the average interest rate difference ($r^{f,*} - r^f$) (or the average slope, $y^{10} - y^{1/4}$), the average foreign bond excess returns for bonds of 30-year maturities in local currency ($rx^{(30),*}$). Panel B reports the average log currency excess returns (rx^{FX}). Panel C reports the average foreign bond excess returns for bonds of 30-year maturities in home currency ($rx^{(30),\$}$). The moments are annualized.

C Structural General Equilibrium Models

This section of the Appendix presents the details of SDF decomposition for three classes of structural general equilibrium models: models with external habit formation, models with long run risks, and models with rare disasters.

C.1 External habit model

The equilibrium log risk-free rate is

$$r_t^f = -E_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) - L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = -\log\delta + \gamma g + \gamma(1-\phi)(\bar{s}-s_t) - \frac{1}{2}\gamma^2\sigma^2(1+\lambda(s_t))^2,$$

which can be also written as

$$r_t^f = -\log\delta + \gamma g + \frac{1}{2}\left(B - \gamma(1-\phi)\right) + B(\bar{s}-s_t) = -\log\delta + \gamma g - \frac{1}{2}\frac{\gamma^2\sigma^2}{\bar{S}^2} + B(\bar{s}-s_t).$$

Therefore, if $B=0$, the log risk-free rate is constant: the intertemporal smoothing effect is exactly offset by the precautionary savings effect. If, on the other hand, $B\neq 0$, then the log risk-free rate is perfectly correlated with the surplus consumption ratio s: it is negatively correlated with s (and hence countercyclical) if $B>0$, and positively correlated with s (and hence procyclical) if $B<0$. This is because, if $B>0$, the intertemporal smoothing effect dominates the precautionary savings effect: when s is above its steady-state level, mean-reversion implies that marginal utility is expected to increase in the future, incentivizing agents to save and decreasing interest rates. On the other hand, if $B<0$, the precautionary savings motive dominates, so agents save more when s is low and marginal utility is more volatile.

To decompose the pricing kernel, we use the guess and verify method. In particular, guess an eigenfunction ϕ of the form

$$\phi(s) = e^{cs},$$

where c is a constant. Then, the (one-period) eigenfunction problem can be written as

$$E_t\left[\exp(\log\delta - \gamma\left(g + (\phi-1)(s_t-\bar{s}) + (1+\lambda(s_t))\varepsilon_{t+1}\right) + cs_{t+1})\right] = \exp(\beta + cs_t)$$

which, after some algebra, yields

$$\log\delta - \gamma g - \gamma(\phi-1)(s_t-\bar{s}) + c(1-\phi)\bar{s} + c\phi s_t + \frac{\sigma^2}{2}\left((c-\gamma)(1+\lambda(s_t)) - c\right)^2 = \beta + cs_t.$$

Setting $c=\gamma$, the expression above becomes

$$\log\delta - \gamma g - \gamma(\phi-1)(s_t-\bar{s}) + \gamma(1-\phi)\bar{s} + \gamma\phi s_t + \frac{\gamma^2\sigma^2}{2} = \beta + \gamma s_t,$$

and, matching the constant terms, we get

$$\beta = \log\delta - \gamma g + \frac{\gamma^2\sigma^2}{2}.$$

Therefore, the transitory component of the pricing kernel is

$$\Lambda_t^{\mathbb{T}} = e^{\beta t - cs_t},$$

so the transitory SDF component is

$$\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}} = e^{\beta - c(s_{t+1}-s_t)} = e^{\log\delta - \gamma g + \frac{\gamma^2\sigma^2}{2} - \gamma\left((1-\phi)(\bar{s}-s_t) + \lambda(s_t)\varepsilon_{t+1}\right)},$$

and the permanent SDF component is

$$\begin{aligned}\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}} = \frac{\Lambda_{t+1}}{\Lambda_t}\left(\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}\right)^{-1} &= e^{\log\delta - \gamma\left(g + (\phi-1)(s_t-\bar{s}) + (1+\lambda(s_t))\varepsilon_{t+1}\right)} e^{-\log\delta + \gamma g - \frac{\gamma^2\sigma^2}{2} + \gamma\left((1-\phi)(\bar{s}-s_t) + \lambda(s_t)\varepsilon_{t+1}\right)} \\ &= e^{-\frac{\gamma^2\sigma^2}{2} - \gamma\varepsilon_{t+1}}.\end{aligned}$$

In the Campbell and Cochrane (1999) model, the permanent SDF component reflects innovations in consumption growth, which permanently affect the level of consumption, whereas the transitory SDF component is driven by innovations in the surplus consumption ratio, which is a stationary variable. However, the two types of innovations are perfectly correlated by assumption, so the two SDF components exhibit positive comovement: a negative consumption growth innovation not only permanently reduces the level of consumption, but also transitorily decreases the surplus consumption ratio of the agent, increasing the local curvature of her utility function. As a result, a negative consumption growth shock implies a positive shock for both SDF components.

Finally, we consider the properties of the SDF and its components. In each country, the conditional entropy of the SDF is

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{\gamma^2\sigma^2}{2}(1+\lambda(s_t))^2 = \frac{\gamma^2\sigma^2}{2}\frac{1}{\bar{S}^2}(1-2(s_t-\bar{s})),$$

the conditional entropy of the permanent SDF component is

$$L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) = \frac{\gamma^2\sigma^2}{2},$$

and the conditional entropy of the transitory SDF component is

$$L_t\left(\frac{\Lambda^{\mathbb{T}}_{t+1}}{\Lambda^{\mathbb{T}}_t}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda^{\mathbb{T}}_{t+1}}{\Lambda^{\mathbb{T}}_t}\right) = \frac{\gamma^2\sigma^2}{2}\lambda(s_t)^2.$$

Notably, the permanent SDF component has constant conditional entropy, whereas the conditional entropy of both the SDF and the transitory SDF component are time varying, as they are functions of the log surplus consumption ratio s. It follows that the conditional term premium, in local currency terms, is

$$E_t\left[rx^{(\infty)}_{t+1}\right] = L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) - L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) = \frac{\gamma^2\sigma^2}{2}\frac{1}{\bar{S}^2}(1-2(s_t-\bar{s})) - \frac{\gamma^2\sigma^2}{2} = \frac{\gamma^2\sigma^2}{2}\left[\frac{1}{\bar{S}^2}(1-2(s_t-\bar{s}))-1\right].$$

C.2 Long-Run Risks Model

We consider two versions of the model: a homeskedastic version, in which $u_t = \theta^u$ and $w_t = \theta^w$ for all t, and the full heteroskedastic version presented in the main text.

Homoskedastic Version We start with the homoskedastic version. We can show that the log SDF has a law of motion given by

$$\log\frac{\Lambda_{t+1}}{\Lambda_t} = A_0 + A_1 x_t + B_1\sqrt{\theta^u}\varepsilon^c_{t+1} + B_2\sqrt{\theta^w}\varepsilon^x_{t+1},$$

where $\{A_0, A_1, B_1, B_2\}$ are constants, the values of which are reported in Panel A of Table 5. If the agent has preferences for early resolution of uncertainty (in which case $\gamma > \frac{1}{\psi}$), then $B_2 < 0$, so a positive shock in the predictable component of consumption growth decreases marginal utility. On the other hand, if the agent prefers early resolution of uncertainty ($\gamma < \frac{1}{\psi}$), then $B_2 > 0$.

The equilibrium log risk-free rate is

$$r^f_t = -E_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) - L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = -A_0 - A_1 x_t - \frac{1}{2}(B_1^2\theta^u + B_2^2\theta^w),$$

so it is positively associated with x, the predictable component of consumption growth: when the drift of consumption growth is above its unconditional mean, the risk-free rate is also above its time-series average.

In order to decompose the SDF, we use the guess and verify method. We guess an eigenfunction ϕ of the form

$$\phi(x,u,w) = e^{cx}$$

where c is a constant. Then, the (one-period) eigenfunction problem can be written as

$$E_t\left[\exp(\log\frac{\Lambda_{t+1}}{\Lambda_t} + cx_{t+1})\right] = \exp(\beta + cx_t)$$

which, exploiting the log-normality of the term inside the expectation, implies

$$E_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t} + cx_{t+1}\right) + \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t} + cx_{t+1}\right) = \beta + cx_t.$$

After some algebra, matching terms yields

$$\beta = A_0 + \frac{1}{2}{B_1}^2\theta^u + \frac{1}{2}(B_2+c)^2\theta^w,$$
$$c = \frac{A_1}{1-\phi^x} = -\frac{1}{\psi}\frac{1}{1-\phi^x} < 0.$$

The transitory component of the pricing kernel is

$$\Lambda_t^{\mathbb{T}} = e^{\beta t - c x_t}$$

so the transitory SDF component is

$$\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}} = e^{\beta + c(1-\phi^x)x_t - c\sqrt{\theta^w}\varepsilon_{t+1}^x} = e^{\beta + A_1 x_t - c\sqrt{\theta^w}\varepsilon_{t+1}^x}$$

and the permanent SDF component is

$$\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}} = \frac{\Lambda_{t+1}}{\Lambda_t}\left(\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}\right)^{-1} = e^{A_0 + A_1 x_t + B_1\sqrt{\theta^u}\varepsilon_{t+1}^c + B_2\sqrt{\theta^w}\varepsilon_{t+1}^x} e^{-\beta - A_1 x_t + c\sqrt{\theta^w}\varepsilon_{t+1}^x},$$

or, after some algebra,

$$\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}} = e^{-\frac{1}{2}B_1{}^2\theta^u - \frac{1}{2}(B_2+c)^2\theta^w + B_1\sqrt{\theta^u}\varepsilon_{t+1}^c + (B_2+c)\sqrt{\theta^w}\varepsilon_{t+1}^x}.$$

Both SDF components are exposed to the consumption drift innovation ε^x, but only the permanent SDF component is exposed to the consumption growth innovation ε^c. As a result, overall SDF and the permanent SDF component have identical loadings on the consumption growth shock. However, the dependence on the consumption drift shock depends on the agent's preferences towards uncertainty resolution. If the agent prefers early resolution of uncertainty ($\gamma > \frac{1}{\psi}$), then a negative consumption drift shock ($\varepsilon^x < 0$) is associated with an increase of the agent's overall SDF and its permanent component and a decline of its transitory component: the loading of the overall log SDF on ε^x is $B_2\sqrt{\theta^w} < 0$, the loading of the permanent SDF component is $(B_2+c)\sqrt{\theta^w} < B_2\sqrt{\theta^w} < 0$ and the loading of the log transitory SDF component is $-c\sqrt{\theta^w} > 0$. This is because the long-run effect of a consumption drift innovation in the pricing kernel (captured by the permanent SDF component) is higher than its short-run effect (captured by the overall SDF). Intuitively, a negative consumption drift shock lowers marginal utility in the long run both through an immediate decline in the continuation utility (reflected in the overall SDF) and through the cumulative effect of a persistent reduction in x, which is equal to $-\frac{1}{\psi}\sum_{j=0}^{\infty}(\phi^x)^j\sqrt{\theta^w} = -\frac{1}{\psi}\frac{1}{1-\phi^x}\sqrt{\theta^w} = c\sqrt{\theta^w}$.

Conditional SDF entropy is constant, given by

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{1}{2}\left(B_1{}^2\theta^u + B_2{}^2\theta^w\right).$$

The conditional entropy of the permanent SDF component is also constant, and equal to

$$L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) = \frac{1}{2}\left(B_1{}^2\theta^u + (B_2+c)^2\theta^w\right),$$

and the conditional entropy of the transitory SDF component is

$$L_t\left(\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}\right) = \frac{1}{2}c^2\theta^w.$$

Notably, the transitory and permanent SDF components are exposed to the expected consumption growth innovation ε^x, but only the permanent SDF component is exposed to the consumption growth innovation ε^c. As a result, the overall SDF and the the permanent SDF component have identical loadings on the consumption growth shock. However, the dependence on the consumption drift shock depends on the agents' attitude towards the resolution of uncertainty

The conditional term premium, in local currency terms, is constant, and equal to

$$E_t\left[rx_{t+1}^{(\infty)}\right] = L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) - L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right) = -\frac{c}{2}\left(c + 2B_2\right),$$

so if the agent prefers early resolution of uncertainty ($\gamma > \frac{1}{\psi}$, so $B_2 < 0$), the conditional term premium is negative. This is because negative consumption drift shocks increase long-run marginal utility more than they increase short-run marginal utility, so long-term bonds hedge long-run risk, as their price increases upon realization of negative consumption drift shocks.

Heteroskedastic Version In this full version of the model, the log SDF follows the law of motion:

$$\log\frac{\Lambda_{t+1}}{\Lambda_t} = A_0 + A_1 x_t + A_2 u_t + A_3 w_t + B_1\sqrt{u_t}\varepsilon_{t+1}^c + B_2\sqrt{w_t}\varepsilon_{t+1}^x + B_3\varepsilon_{t+1}^u + B_4\varepsilon_{t+1}^w,$$

Table 5: Pricing Kernel Loadings in the Long Run Risks Model

Loadings	Parameters	Loadings	Parameters
			Panel A: Homoskedastic Model $\log \frac{\Lambda_{t+1}}{\Lambda_t} = A_0 + A_1 x_t + B_1\sqrt{\theta^u}\varepsilon^c_{t+1} + B_2\sqrt{\theta^w}\varepsilon^x_{t+1}$.
A_1	$-\frac{1}{\psi}$	B_1	$-\gamma$
		B_2	$\left(\frac{1}{\psi}-\gamma\right)\frac{\kappa}{1-\kappa\phi^x}$
			Transitory Component $\log \Lambda^{\mathbb{T}}_t = \beta t - c x_t$
c_1	$\frac{A_1}{1-\phi^x}$	β	$A_0 + \frac{1}{2}{B_1}^2\theta^u + \frac{1}{2}(B_2+c)^2\theta^w$,
			Panel B: Heteroskedastic Model $\log \frac{\Lambda_{t+1}}{\Lambda_t} = A_0 + A_1 x_t + A_2 u_t + A_3 w_t + B_1\sqrt{u_t}\varepsilon^c_{t+1} + B_2\sqrt{w_t}\varepsilon^x_{t+1} + B_3\varepsilon^u_{t+1} + B_4\varepsilon^w_{t+1}$.
A_1	$-\frac{1}{\psi}$	B_1	$-\gamma$
A_2	$\left(\frac{1}{\psi}-\gamma\right)\frac{\gamma-1}{2}$	B_2	$\left(\frac{1}{\psi}-\gamma\right)\frac{\kappa}{1-\kappa\phi^x}$
A_3	$\left(\frac{1}{\psi}-\gamma\right)\frac{\gamma-1}{2}\left(\frac{\kappa}{1-\kappa\phi^x}\right)^2$	B_3	$\left(\frac{1}{\psi}-\gamma\right)\frac{1-\gamma}{2}\frac{\kappa}{1-\kappa\phi^u}\sigma^u$
		B_4	$\left(\frac{1}{\psi}-\gamma\right)\left(\frac{\kappa}{1-\kappa\phi^x}\right)^2\frac{\kappa}{1-\kappa\phi^w}\sigma^w$.
			Transitory Component $\log \Lambda^{\mathbb{T}}_t = \beta t - c_1 x_t - c_2 u_t - c_3 w_t$
c_1	$\frac{A_1}{1-\phi^x}$	c_2	$\frac{A_2+\frac{1}{2}B_1^2}{1-\phi^u}$
c_3	$\frac{A_3+\frac{1}{2}(B_2+\frac{A_1}{1-\phi^x})^2}{1-\phi^w}$	β	$A_0 + c_2(1-\phi^u)\theta^u + c_3(1-\phi^w)\theta^w + \frac{1}{2}(B_3+c_2\sigma^u)^2 + \frac{1}{2}(B_4+c_3\sigma^w)^2$.

Notes: Pricing kernel loading parameters in the long run risks model. Parameter κ is defined as $\kappa \equiv \frac{\delta e^{\left(1-\frac{1}{\psi}\right)\bar{m}}}{1-\delta+\delta e^{\left(1-\frac{1}{\psi}\right)\bar{m}}}$, where $\bar{m}$ is the point around which a log-linear approximation is taken (see Engel (2016) for details); if $\bar{m}=0$, then $\kappa=\delta$.

where $\{A_0, A_1, A_2, A_3, B_1, B_2, B_3, B_4\}$ are constants, the values of which are reported in Panel B of Table 5. As usual, we assume that the agent has preferences for early resolution of uncertainty ($\gamma > \frac{1}{\psi}$), so $B_2 < 0$.

The equilibrium log risk-free rate is

$$r_t^f = -E_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) - L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = -A_0 - A_1 x_t - A_2 u_t - A_3 w_t - \frac{1}{2}\left({B_1}^2 u_t + {B_2}^2 w_t + {B_3}^2 + {B_4}^2\right),$$

or

$$r_t^f = -A_0 - \frac{1}{2}\left({B_3}^2 + {B_4}^2\right) - A_1 x_t - \left(A_2 + \frac{{B_1}^2}{2}\right)u_t - \left(A_3 + \frac{{B_2}^2}{2}\right)w_t,$$

or

$$r_t^f = -A_0 - \frac{1}{2}\left({B_3}^2 + {B_4}^2\right) + \frac{1}{\psi}x_t - \frac{1}{2}\left(\frac{\gamma-1}{\psi} + \gamma\right)u_t - \frac{1}{2}\left(\frac{1}{\psi} - \gamma\right)\left(\frac{1}{\psi} - 1\right)\left(\frac{\kappa}{1-\kappa\phi^x}\right)^2 w_t.$$

Thus, the risk-free rate is positively associated with x, the predictable component of consumption growth, due to the intertemporal smoothing effect, and negatively associated with u, the conditional variance of the consumption growth shock, as the intertemporal smoothing effect is dominated by the precautionary savings effect. Finally, the sign of the relationship between the risk-free rate and w, the conditional variance of the consumption drift shock, depends on the value of the IES parameter: if $\psi > 1$, then the relationship is negative, as the precautionary savings effect dominates, whereas if $\psi < 1$, then the relationship is positive, as the intertemporal smoothing effect dominates.

To decompose the pricing kernel, we use the guess and verify method. In particular, guess an eigenfunction ϕ of the form

$$\phi(x, u, w) = e^{c_1 x + c_2 u + c_3 w}$$

where $\{c_1, c_2, c_3\}$ are constants. Then, the (one-period) eigenfunction problem can be written as

$$E_t\left[\exp(\log\frac{\Lambda_{t+1}}{\Lambda_t} + c_1 x_{t+1} + c_2 u_{t+1} + c_3 w_{t+1})\right] = \exp(\beta + c_1 x_t + c_2 u_t + c_3 w_t)$$

which, exploiting the log-normality of the term inside the expectation, implies

$$E_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t} + c_1 x_{t+1} + c_2 u_{t+1} + c_3 w_{t+1}\right) + \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t} + c_1 x_{t+1} + c_2 u_{t+1} + c_3 w_{t+1}\right) = \beta + c_1 x_t + c_2 u_t + c_3 w_t.$$

After some algebra, matching terms yields

$$\beta = A_0 + c_2(1-\phi^u)\theta^u + c_3(1-\phi^w)\theta^w + \frac{1}{2}(B_3 + c_2\sigma^u)^2 + \frac{1}{2}(B_4 + c_3\sigma^w)^2$$

$$c_1 = A_1 + c_1\phi^x$$

$$c_2 = A_2 + c_2\phi^u + \frac{1}{2}B_1^2$$

$$c_3 = A_3 + c_3\phi^w + \frac{1}{2}(B_2 + c_1)^2$$

so

$$c_1 = \frac{A_1}{1-\phi^x} = -\frac{1}{\psi}\frac{1}{1-\phi^x} < 0,$$

$$c_2 = \frac{A_2 + \frac{1}{2}B_1^2}{1-\phi^u} = \frac{1}{2}\left(\frac{\gamma-1}{\psi} + \gamma\right)\frac{1}{1-\phi^u} > 0,$$

$$c_3 = \frac{A_3 + \frac{1}{2}(B_2 + c_1)^2}{1-\phi^w} = \frac{\left(\frac{1}{\psi} - \gamma\right)\frac{\gamma-1}{2}\left(\frac{\kappa}{1-\kappa\phi^x}\right)^2 + \frac{1}{2}\left(\left(\frac{1}{\psi} - \gamma\right)\frac{\kappa}{1-\kappa\phi^x} - \frac{1}{\psi}\frac{1}{1-\phi^x}\right)^2}{1-\phi^w} > 0,$$

where the sign for c_2 and c_3 is determined under the assumption that $\gamma > \frac{1}{\psi}$. The transitory component of the pricing kernel is

$$\Lambda_t^{\mathbb{T}} = e^{\beta t - c_1 x_t - c_2 u_t - c_3 w_t},$$

so the transitory SDF component is

$$\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}} = e^{\beta + c_1(1-\phi^x)x_t - c_2(1-\phi^u)(\theta^u - u_t) - c_3(1-\phi^w)(\theta^w - w_t) - c_1\sqrt{w_t}\varepsilon_{t+1}^x - c_2\sigma^u\varepsilon_{t+1}^u - c_3\sigma^w\varepsilon_{t+1}^w},$$

and the permanent SDF component is

$$\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t} = \frac{\Lambda_{t+1}}{\Lambda_t}\left(\frac{\Lambda^{\mathbb{T}}_{t+1}}{\Lambda^{\mathbb{T}}_t}\right)^{-1} = e^{A_0+A_1x_t+A_2u_t+A_3w_t+B_1\sqrt{u_t}\varepsilon^c_{t+1}+B_2\sqrt{w_t}\varepsilon^x_{t+1}+B_3\varepsilon^u_{t+1}+B_4\varepsilon^w_{t+1}} \times$$
$$e^{-\beta-c_1(1-\phi^x)x_t+c_2(1-\phi^u)(\theta^u-u_t)+c_3(1-\phi^w)(\theta^w-w_t)+c_1\sqrt{w_t}\varepsilon^x_{t+1}+c_2\sigma^u\varepsilon^u_{t+1}+c_3\sigma^w\varepsilon^w_{t+1}},$$

or, after some algebra,

$$\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t} = e^{-(1/2)(B_3+c_2\sigma^u)^2-(1/2)(B_4+c_3\sigma_w)^2-(1/2)B_1^2u_t-(1/2)(B_2+c_1)^2w_t} \times$$
$$e^{B_1\sqrt{u_t}\varepsilon^c_{t+1}+(B_2+c_1)\sqrt{w_t}\varepsilon^x_{t+1}+(B_3+c_2\sigma^u)\varepsilon^u_{t+1}+(B_4+c_3\sigma^w)\varepsilon^w_{t+1}}.$$

In summary, both SDF components are exposed to the consumption drift innovation ε^x, the consumption growth variance innovation ε^u, and the consumption drift variance innovation ε^w, but only the permanent SDF component is exposed to the consumption growth innovation ε^c. As a result, overall SDF and the permanent SDF component have identical loadings on the consumption growth shock. However, the dependence on the rest of the innnovations depends on the agent's preferences regarding the resolutions of uncertainty. If the agent prefers early resolution ($\gamma > \frac{1}{\psi}$), we have $c_1 < 0$, $c_2 > 0$ and $c_3 > 0$.

We can start with exposure to consumption drift shocks. Since $B_2 < 0$, $c_1 < 0$ implies that the permanent SDF component is more sensitive to consumption drift shocks than the total SDF, while the transitory SDF component has the opposite sign. For example, a negative consumption drift shock ($\varepsilon^x < 0$) is associated with an increase of the agent's overall SDF and its permanent component and a decline of its transitory component. This is because the long-run effect of a consumption drift innovation in the pricing kernel (captured by the permanent SDF component) is higher than its short-run effect (captured by the overall SDF). Intuitively, a negative consumption drift shock lowers marginal utility in the long run both through an immediate decline in the continuation utility (reflected in the overall SDF) and through the cumulative effect of a persistent reduction in x, which is equal to $-\frac{1}{\psi}\sum_{j=0}^{\infty}(\phi^x)^j\sqrt{\theta^w} = -\frac{1}{\psi}\frac{1}{1-\phi^x}\sqrt{\theta^w} = c_1\sqrt{\theta^w}$.

As regards the two variance shocks, whether long-run marginal utility reacts more or less than short-run marginal utility depends on the sign of B_3 and B_4. If $\gamma > 1$, i.e. the agent is more risk-averse than a log utility investor, then $B_3 > 0$ and $B_4 > 0$, so short-run marginal utility increases upon realization of any positive variance shock. Thus, $c_2 > 0$ and $c_3 > 0$ imply that long-run marginal utility reacts more than short-run marginal utility: when either $\varepsilon^u > 0$ or $\varepsilon^w > 0$, the permanent SDF component increases more than total SDF, with the transitory SDF component declining. On the other hand, if $\gamma < 1$, then $B_3 < 0$ and $B_4 < 0$, in which case short-run marginal utility declines upon realization of any positive variance shock. As a result, $c_2 > 0$ and $c_3 > 0$ imply that long-run marginal utility reacts less than short-run marginal utility: when either $\varepsilon^u > 0$ or $\varepsilon^w > 0$, the permanent SDF component declines less than total SDF, as the transitory SDF component also falls.

Conditional SDF entropy is

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right) = \frac{1}{2}\left({B_1}^2u_t + {B_2}^2w_t + {B_3}^2 + {B_4}^2\right),$$

whereas the conditional entropy of the permanent SDF component is

$$L_t\left(\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) = \frac{1}{2}var_t\left(\log\frac{\Lambda^{\mathbb{P}}_{t+1}}{\Lambda^{\mathbb{P}}_t}\right) = \frac{1}{2}\left({B_1}^2u_t + (B_2+c_1)^2w_t + (B_3+c_2\sigma^u)^2 + (B_4+c_3\sigma^w)^2\right).$$

For conditional SDF entropy to be identical across countries, it is sufficient that the conditional variances u and w are identical across countries, that $B_1 = B_1^*$ and $B_2 = B_2^*$ (i.e. that $\gamma = \gamma^*$ and $\left(\frac{1}{\psi} - \gamma\right)\frac{\delta}{1-\delta\phi^x} = \left(\frac{1}{\psi^*} - \gamma^*\right)\frac{\delta^*}{1-\delta^*\phi^{x,*}}$), and that $(B_3+c_2\sigma^u)^2 + (B_4+c_3\sigma^w)^2 = (B_3^*+c_2^*\sigma^{u,*})^2 + (B_4^*+c_3^*\sigma^{w,*})^2$. For the conditional entropy of the permanent SDF component to be identical across countries, we need $B_2 + c_1 = B_2^* + c_1^*$ instead of $B_2 = B_2^*$. Therefore, we will have non-identical SDF entropy and identical entropy of the permanent SDF component across countries if $B_2 + c_1 = B_2^* + c_1^*$ and $c_1 - c_1^* = B_2^* - B_2 \neq 0$. For example, those conditions are satisfied if $\gamma = \gamma^*$, $\delta = \delta^*$ and $\psi = \psi^*$, but $\phi^x \neq \phi^{x,*}$ such that $(1-\delta\phi^x)(1-\delta\phi^{x,*}) = \delta^2(1-\gamma\psi)(1-\phi^x)(1-\phi^{x,*})$.

Finally, the term premium, in local currency terms, is

$$E_t\left[rx^{(\infty)}_{t+1}\right] = \frac{1}{2}\left({B_2}^2 - (B_2+c_1)^2\right)w_t + \frac{1}{2}\left({B_3}^2 - (B_3+c_2\sigma^u)^2\right) + \frac{1}{2}\left({B_4}^2 - (B_4+c_3\sigma^w)^2\right).$$

Following the discussion above, if $\gamma > \frac{1}{\psi}$ (in which case $B_2 < 0$), then the conditional term premium is negatively associated with w, the variance of the consumption growth drift. This is because negative consumption drift shocks increase long-run marginal utility more than they increase short-run marginal utility, so long-term bonds hedge long-run risk, as their price increases upon realization of negative consumption drift shocks. Therefore, the higher the conditional volatility of those shocks, the more attractive long-term bonds are as a hedging asset, and the lower risk premium they earn.

C.3 Disasters Model

In the Farhi and Gabaix (2016) rare disasters model, the SDF has law of motion

$$\frac{\Lambda_{t+1}}{\Lambda_t} = \frac{\Lambda^*_{t+1}}{\Lambda^*_t}\frac{\omega_{t+1}}{\omega_t}\frac{1+Ax_{t+1}}{1+Ax_t},$$

where

$$\frac{\Lambda^*_{t+1}}{\Lambda^*_t} = e^{-R} \times \begin{cases} 1, \text{if there no disaster at } t+1 \\ B^{-\gamma}_{t+1}, \text{if there is a disaster at } t+1 \end{cases}$$

is the global numeraire SDF,

$$\frac{\omega_{t+1}}{\omega_t} = e^{g_\omega} \times \begin{cases} 1, \text{if there is no disaster at } t+1 \\ F_{t+1}, \text{if there is a disaster at } t+1 \end{cases}$$

is the productivity growth of the country, and x is defined as $x_t \equiv e^{-h_*}\hat{H}_t$, where $\hat{H}$ is the time-varying component of the resilience of the country, to be discussed below. Finally, $A \equiv \frac{e^{-R-\lambda+g_\omega+h_*}}{1-e^{-R-\lambda+g_\omega+h_*-\phi_H}}$, where λ is the investment depreciation rate, and $h_* \equiv \log(1+H_*)$. Finally, we assume that $R+\lambda-g_\omega-h_*>0$, so $A>0$.

Resilience is defined as

$$H_t = H_* + \hat{H}_t = p_t E^D_t\left[B^\gamma_{t+1}F_{t+1}-1\right],$$

where p_t is the conditional probability of a disaster occurring next period and E^D_t is the period t expectation conditional on a disaster occurring next period. The time-varying component of resilience has law of motion

$$\hat{H}_{t+1} = \frac{1+H_*}{1+H_t}e^{-\phi_H}\hat{H}_t + \varepsilon^H_{t+1},$$

with the conditional expectation of ε^H being zero independently of the realization of a disaster. As a result, the conditional expectation of x is

$$E_t(x_{t+1}) = e^{-\phi_H}\frac{x_t}{1+x_t}.$$

The equilibrium log risk-free rate is

$$r^f_t = -\log E_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right) = (R-g_\omega-h_*) + \log\left(\frac{1+Ax_t}{1+(Ae^{-\phi_H}+1)x_t}\right),$$

so it is decreasing in x.

To decompose the pricing kernel, we use the guess and verify method. In particular, guess an eigenfunction ϕ of the form

$$\phi(x) = \frac{c+x}{1+Ax},$$

where c is a constant.[20] Then, the (one-period) eigenfunction problem can be written as

$$E_t\left[\frac{\Lambda^*_{t+1}}{\Lambda^*_t}\frac{\omega_{t+1}}{\omega_t}\frac{1+Ax_{t+1}}{1+Ax_t}\frac{c+x_{t+1}}{1+Ax_{t+1}}\right] = e^\beta\frac{c+x_t}{1+Ax_t}$$

which yields

$$E_t\left[\frac{\Lambda^*_{t+1}}{\Lambda^*_t}\frac{\omega_{t+1}}{\omega_t}\right]E_t\left[\frac{1+Ax_{t+1}}{1+Ax_t}\frac{c+x_{t+1}}{1+Ax_{t+1}}\right] = e^\beta\frac{c+x_t}{1+Ax_t}$$

so

$$e^{-R+g_\omega+h_*}(1+x_t)E_t\left[\frac{1+Ax_{t+1}}{1+Ax_t}\frac{c+x_{t+1}}{1+Ax_{t+1}}\right] = e^\beta\frac{c+x_t}{1+Ax_t}.$$

The expression above becomes:

$$e^{-R+g_\omega+h_*}(1+x_t)E_t\left[c+x_{t+1}\right] = e^\beta(c+x_t),$$

[20] It can be shown that conjecturing the more general eigenfunction $\phi(x) = \frac{c_1+c_2x}{1+Ax}$, where c_1 and c_2 are constants, leads to same SDF decomposition as the one derived below.

so, plugging in the expression for the conditional expectation of x, we get

$$e^{-R+g_\omega+h_*}(1+x_t)\left(c+e^{-\phi_H}\frac{x_t}{1+x_t}\right)=e^{\beta}(c+x_t),$$

which yields

$$\beta=-R+g_\omega+h_*$$

and

$$c=1-e^{-\phi_H}.$$

The lower bound of x is $e^{-\phi_H}-1$, so $c+x_t>0$ for all t; thus, the conjectured eigenfunction is strictly positive, as required. The transitory component of the pricing kernel is

$$\Lambda_t^{\mathbb{T}}=e^{\beta t}\frac{1+Ax_t}{c+x_t}$$

so the transitory SDF component is

$$\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}=e^{\beta}\frac{1+Ax_{t+1}}{c+x_{t+1}}\frac{c+x_t}{1+Ax_t}=e^{-R+g_\omega+h_*}\frac{1+Ax_{t+1}}{1+Ax_t}\frac{c+x_t}{c+x_{t+1}}$$

and the permanent SDF component is

$$\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}=\frac{\Lambda_{t+1}}{\Lambda_t}\left(\frac{\Lambda_{t+1}^{\mathbb{T}}}{\Lambda_t^{\mathbb{T}}}\right)^{-1}\quad=\quad e^{R-g_\omega-h_*}\frac{\Lambda_{t+1}^{*}}{\Lambda_t^{*}}\frac{\omega_{t+1}}{\omega_t}\frac{c+x_{t+1}}{c+x_t}.$$

The transitory SDF component is only exposed to resilience shocks (ε^H), but not to disaster risk; the entirety of the disaster risk for marginal utility is reflected in the permanent SDF component, as disasters permanently affect the future level of marginal utility.

We can now calculate the conditional entropy of the SDF and its components. It holds that

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right)=\log E_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right)-E_t\left(\log\frac{\Lambda_{t+1}}{\Lambda_t}\right),$$

so we can write

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right)=L_t\left(\frac{\Lambda_{t+1}^{*}}{\Lambda_t^{*}}\frac{\omega_{t+1}}{\omega_t}\right)+L_t\left(\frac{1+Ax_{t+1}}{1+Ax_t}\right).$$

After some algebra, we get

$$L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right)=\log(1+H_t)-p_tE_t^D\left[\log(B_{t+1}^{\gamma}F_{t+1})\right]+L_t\left(\frac{1+Ax_{t+1}}{1+Ax_t}\right).$$

Similarly, the conditional entropy of the permanent SDF component is

$$L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right)=\log(1+H_t)-p_tE_t^D\left[\log(B_{t+1}^{\gamma}F_{t+1})\right]+L_t\left(\frac{c+x_{t+1}}{c+x_t}\right).$$

Therefore, the conditional term premium, in local currency terms, is

$$E_t\left[rx_{t+1}^{(\infty)}\right]=L_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\right)-L_t\left(\frac{\Lambda_{t+1}^{\mathbb{P}}}{\Lambda_t^{\mathbb{P}}}\right)=L_t\left(\frac{1+Ax_{t+1}}{1+Ax_t}\right)-L_t\left(\frac{c+x_{t+1}}{c+x_t}\right).$$